Preface
Biology is designed for multi-semester biology courses for science majors. It is grounded on an evolutionary basis and includes exciting features that highlight careers in the biological sciences and everyday applications of the concepts at hand. To meet the needs of today's instructors and students, some content has been strategically condensed while maintaining the overall scope and coverage of traditional texts for this course. Instructors can customize the book, adapting it to the approach that works best in their classroom. Biology also includes an innovative art program that incorporates critical thinking and clicker questions to help students understand—and apply—key concepts.

Welcome to *Biology 2e* (2nd edition), an OpenStax resource. This textbook was written to increase student access to high-quality learning materials, maintaining highest standards of academic rigor at little to no cost.

About OpenStax

OpenStax is a nonprofit based at Rice University, and it's our mission to improve student access to education. Our first openly licensed college textbook was published in 2012, and our library has since scaled to over 25 books for college and AP® courses

used by hundreds of thousands of students. OpenStax Tutor, our low-cost personalized learning tool, is being used in college courses throughout the country. Through our partnerships with philanthropic foundations and our alliance with other educational resource organizations, OpenStax is breaking down the most common barriers to learning and empowering students and instructors to succeed.

About OpenStax resources

Customization

Biology 2e is licensed under a Creative Commons Attribution 4.0 International (CC BY) license, which means that you can distribute, remix, and build upon the content, as long as you provide attribution to OpenStax and its content contributors.

Because our books are openly licensed, you are free to use the entire book or pick and choose the sections that are most relevant to the needs of your course. Feel free to remix the content by assigning your students certain chapters and sections in your syllabus, in the order that you prefer. You can even provide a direct link in your syllabus to the sections in the web view of your book.

Instructors also have the option of creating a customized version of their OpenStax book. The custom version can be made available to students in low-cost print or digital form through their campus bookstore. Visit the Instructor Resources section of your book page on OpenStax.org for more information.

Errata

All OpenStax textbooks undergo a rigorous review process. However, like any professional-grade textbook, errors sometimes occur. Since our books are web based, we can make updates periodically when deemed pedagogically necessary. If you have a correction to suggest, submit it through the link on your book page on OpenStax.org. Subject matter experts review all errata suggestions. OpenStax is committed to remaining transparent about all updates, so you will also find a list of past errata changes on your book page on OpenStax.org.

Format

You can access this textbook for free in web view or PDF through OpenStax.org, and for a low cost in print.

About *Biology 2e*

Biology 2e (2nd edition) is designed to cover the scope and sequence requirements of a typical two-semester biology course for science majors. The text provides comprehensive coverage of foundational research and core biology concepts through an evolutionary lens. *Biology* includes rich features that engage students in scientific inquiry, highlight careers in the biological sciences, and offer everyday applications. The book also includes various types of practice and homework questions that help students understand — and apply — key concepts.

The 2nd edition has been revised to incorporate clearer, more current, and more dynamic explanations, while maintaining the same organization as the first edition. Art and illustrations have been substantially improved, and the textbook features additional assessments and related resources.

Coverage and scope

Biology was one of the first textbooks published by OpenStax and has been used by hundreds of faculty and thousands of students since 2012. We mined our adopters' extensive and helpful feedback to identify the most significant revision needs while maintaining the organization that many instructors had incorporated into their courses. Specific surveys, focus groups, and pre-revision reviews, as well as data from our OpenStax Tutor users, all

aided in planning the revision.

The result is a book that thoroughly treats biology's foundational concepts while adding current and meaningful coverage in specific areas. *Biology 2e* retains its manageable scope and contains ample features to draw learners into the discipline.

Structurally, the textbook remains similar to the first edition, with no chapter reorganization and very targeted changes at the section level (mostly in biodiversity).

- **Unit 1: The Chemistry of Life.** Our opening unit introduces students to the sciences, including the scientific method and the fundamental concepts of chemistry and physics that provide a framework within which learners comprehend biological processes.
- **Unit 2: The Cell.** Students will gain solid understanding of the structures, functions, and processes of the most basic unit of life: the cell.
- **Unit 3: Genetics.** Our comprehensive genetics unit takes learners from the earliest experiments that revealed the basis of genetics through the intricacies of DNA to current applications in the emerging studies of biotechnology and genomics.
- **Unit 4: Evolutionary Processes.** The core concepts of evolution are discussed in this unit with examples illustrating evolutionary

processes. Additionally, the evolutionary basis of biology reappears throughout the textbook in general discussion and is reinforced through special call-out features highlighting specific evolution-based topics.

Unit 5: Biological Diversity. The diversity of life is explored with detailed study of various organisms and discussion of emerging phylogenetic relationships. This unit moves from viruses to living organisms like bacteria, discusses the organisms formerly grouped as protists, and devotes multiple chapters to plant and animal life.

Unit 6: Plant Structure and Function. Our plant unit thoroughly covers the fundamental knowledge of plant life essential to an introductory biology course.

Unit 7: Animal Structure and Function. An introduction to the form and function of the animal body is followed by chapters on specific body systems and processes. This unit touches on the biology of all organisms while maintaining an engaging focus on human anatomy and physiology that helps students connect to the topics.

Unit 8: Ecology. Ecological concepts are broadly covered in this unit, with features highlighting localized, real-world issues of conservation and biodiversity.

Changes to the Second Edition

OpenStax only undertakes second editions when significant modifications to the text are necessary. In the case of *Biology 2e*, user feedback indicated that we needed to focus on a few key areas, which we have done in the following ways:

- **Content revisions for clarity, accuracy, and currency.** The revision plan varied by chapter based on need. About twenty chapters were wholly revised with significant updates to conceptual coverage, research-informed data, and clearer language. In about fifteen other chapters, the revisions focused mostly on readability and clearer language with fewer conceptual and factual changes.
- **Additional end-of-chapter questions.** The authors added new assessments to nearly every chapter, including both review and critical thinking questions. The additions total over 350 new items.
- **Art and illustrations.** Under the guidance of the authors and expert scientific illustrators, especially those well versed in creating accessible art, the OpenStax team made changes to most of the art in *Biology*. You will find examples in the section below. The revisions fall into the following categories:
  - Revisions for accuracy
  - Redesigns for greater understanding and impact

- Recoloring art for overall consistency

Accessibility improvements. As with all OpenStax books, the first edition of *Biology* was created with a focus on accessibility. We have emphasized and improved that approach in the second edition.

- To accommodate users of specific assistive technologies, all alternative text was reviewed and revised for comprehensiveness and clarity.
- Many illustrations were revised to improve the color contrast, which is important for some visually impaired students.
- Overall, the OpenStax platform has been continually upgraded to improve accessibility.

A transition guide will be available on OpenStax.org to highlight the specific chapter-level changes to the second edition.

Pedagogical foundation

The pedagogical choices, chapter arrangements, and learning objective fulfillment were developed and vetted with the feedback of another one hundred reviewers, who thoroughly read the material and offered detailed critical commentary.

- **Evolution Connection** features uphold the importance of evolution to all biological study through discussions like “The Evolution of Metabolic Pathways” and “Algae and Evolutionary Paths to Photosynthesis.”
- **Scientific Method Connection** call-outs walk students through actual or thought experiments that elucidate the steps of the scientific process as applied to the topic. Features include “Determining the Time Spent in Cell Cycle Stages” and “Testing the Hypothesis of Independent Assortment.”
- **Career Connection** features present information on a variety of careers in the biological sciences, introducing students to the educational requirements and day-to-day work life of a variety of professions, such as microbiologist, ecologist, neurologist, and forensic scientist.
- **Everyday Connection** features tie biological concepts to emerging issues and discuss science in terms of everyday life. Topics include “Chesapeake Bay” and “Can Snail Venom Be Used as a Pharmacological Pain Killer?”

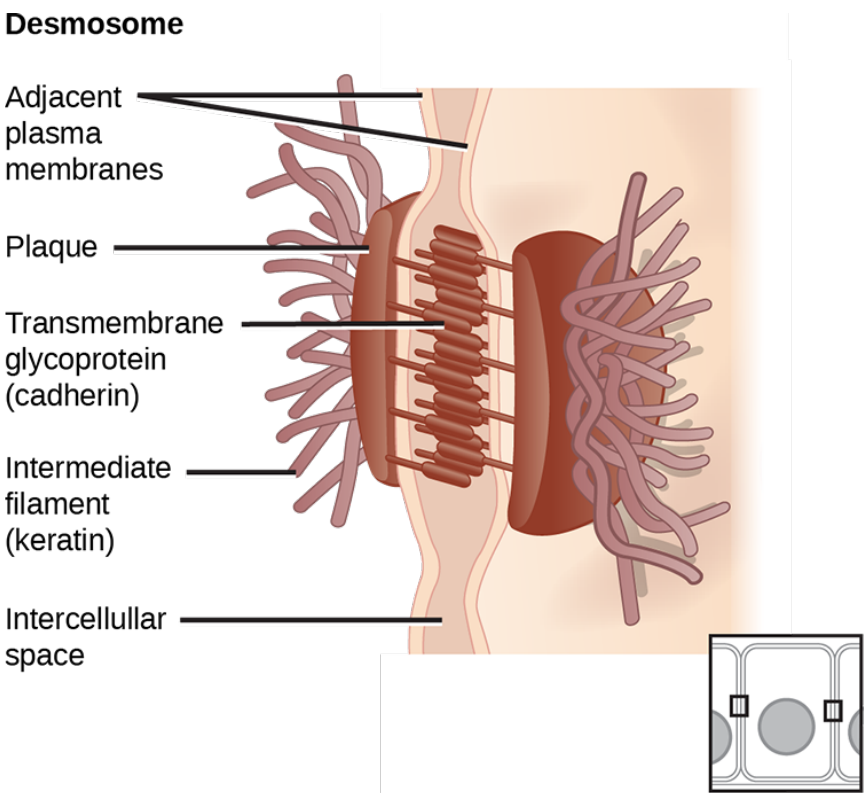
Art and animations that engage

Our art program takes a straightforward approach designed to help students learn the concepts of biology through simple, effective illustrations, photos, and micrographs. Biology 2e also

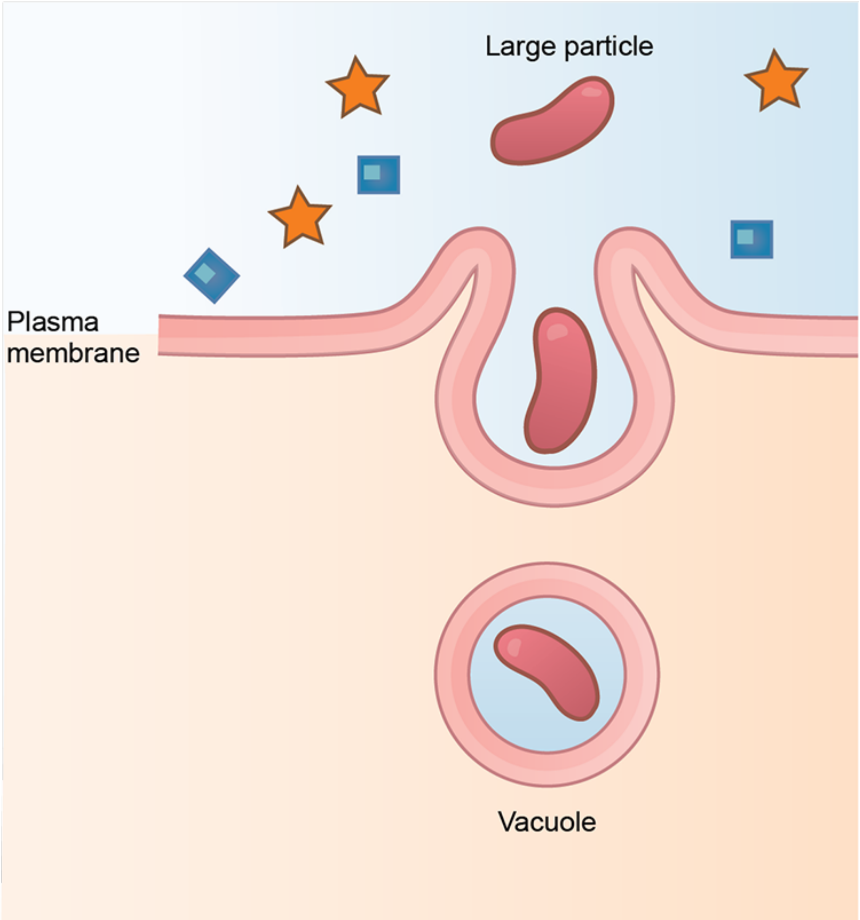
incorporates links to relevant animations and interactive exercises that help bring biology to life for students.

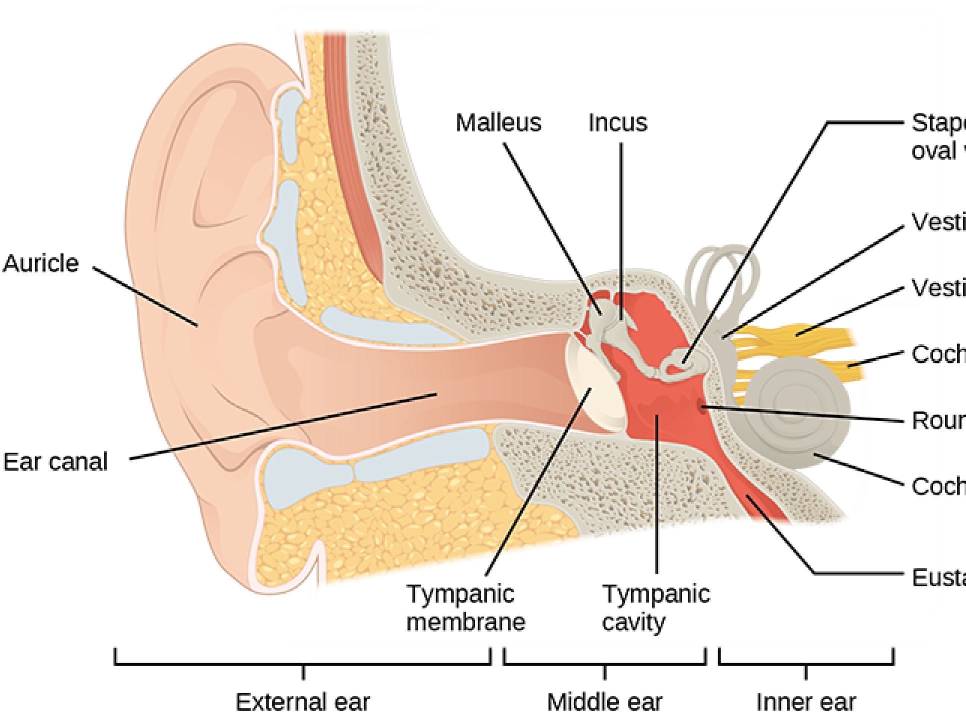
- **Visual Connection** features call out core figures in each chapter for student study. Questions about key figures, including clicker questions that can be used in the classroom, engage students' critical thinking and analytical abilities to ensure their genuine understanding.
- **Link to Learning** features direct students to online interactive exercises and animations to add a fuller context and examples to core content.

Below are a few examples of the revised art for *Biology 2e*:


Desmosome
Adjacent plasma membranes
Plaque
Transmembrane glycoprotein (cadherin)
Intermediate filament (keratin)
Intercellular space

Phagocytosis





Additional resources

Student and instructor resources

We’ve compiled additional resources for both students and instructors, including Getting Started Guides, an instructor solution guide, and PowerPoint lecture slides. Instructor resources require a verified instructor account, which you can apply for when you log in or create your account on OpenStax.org. Take advantage of these resources to supplement your OpenStax book.

Community Hubs

OpenStax partners with the Institute for the Study of Knowledge Management in Education (ISKME) to offer Community Hubs on OER Commons – a platform for instructors to share community-created resources that support OpenStax books, free of charge. Through our Community Hubs, instructors can upload their own materials or download resources to use in their own courses, including additional ancillaries, teaching material, multimedia, and relevant course content. We encourage instructors to join the hubs for the subjects most relevant to your teaching and research as an opportunity both to enrich your courses and to engage with other faculty.

To reach the Community Hubs, visit www.oercommons.org/hubs/OpenStax.

Technology partners

As allies in making high-quality learning materials accessible, our technology partners offer optional low-cost tools that are integrated with OpenStax books. To access the technology options for your text, visit your book page on OpenStax.org.

About the authors

Second edition authors and reviewers

Senior Contributing Authors

Mary Ann Clark, Texas Wesleyan University
Jung Choi, Georgia Institute of Technology
Matthew Douglas, Grand Rapids Community College

Reviewers

Kathleen Berlyn, Baltimore City Community College
Bridgett Brinton, Armstrong State University
Jennifer Chase, Northwest Nazarene University
Amy Hoffman, Grayson County College
Olga Kopp, Utah Valley University
Jennifer Larson, Capital University
Jason Locklin, Austin Community College
Hongmei Ma, American University
Melissa Masse, Tulsa Community College
Shannon McDermott, Central Virginia Community College
Bryan Monesson-Olson, University of Massachusetts Amherst
Amber Reece, California State University Fresno
Monique Reed, Texas A&M University
Jeffrey Roberts, American River College
Matthew Smith, North Dakota State University
Dawn Wankowski, Cardinal Stritch University

First edition authors and reviewers

Senior Contributing Authors

Yael Avissar (Cell Biology), Rhode Island College

Jung Choi (Genetics), Georgia Institute of Technology
Jean DeSaix (Evolution), University of North Carolina at Chapel Hill
Vladimir Jurukovski (Animal Physiology), Suffolk County Community College
Robert Wise (Plant Biology), University of Wisconsin, Oshkosh
Connie Rye (General Content Lead), East Mississippi Community College

Contributing Authors and Reviewers

Julie Adams, Aurora University
Summer Allen, Brown University
James Bader, Case Western Reserve University
David Bailey, St. Norbert College
Mark Belk, Brigham Young University
Nancy Boury, Iowa State University
Lisa Bonneau, Metropolitan Community College – Blue River
Graciela Brelles-Marino, California State University Pomona
Mark Browning, Purdue University
Sue Chaplin, University of St. Thomas
George Cline, Jacksonville State University
Deb Cook, Georgia Gwinnett College
Diane Day, Clayton State University
Frank Dirrigl, The University of Texas Pan American
Waneene Dorsey, Grambling State University
Nick Downey, University of Wisconsin La Crosse
Rick Duhrkopf, Baylor University

Kristy Duran, Adams State University
Stan Eisen, Christian Brothers University
Brent Ewers, University of Wyoming
Myriam Feldman, Lake Washington Institute of Technology
Michael Fine, Virginia Commonwealth University
Linda Flora, Delaware County Community College
Thomas Freeland, Walsh University
David Grisé, Texas A & M University – Corpus Christi
Andrea Hazard, SUNY Cortland
Michael Hedrick, University of North Texas
Linda Hensel, Mercer University
Mark Kopeny, University of Virginia
Norman Johnson, University of Massachusetts Amherst
Grace Lasker, Lake Washington Institute of Technology; Walden University
Sandy Latourelle, SUNY Plattsburgh
Theo Light, Shippensburg University
Clark Lindgren, Grinnell College
James Malcolm, University of Redlands
Mark Meade, Jacksonville State University
Richard Merritt, Houston Community College
James Mickle, North Carolina State University
Jasleen Mishra, Houston Community College
Dudley Moon, Albany College of Pharmacy and Health Sciences
Shobhana Natarajan, Brookhaven College
Jonas Okeagu, Fayetteville State University
Diana Oliveras, University of Colorado Boulder

John Peters, College of Charleston
Joel Piperberg, Millersville University
Johanna Porter-Kelley, Winston-Salem State University
Robyn Puffenbarger, Bridgewater College
Dennis Revie, California Lutheran University
Ann Rushing, Baylor University
Sangha Saha, City College of Chicago
Edward Saiff, Ramapo College of New Jersey
Brian Shmaefsky, Lone Star College System
Robert Sizemore, Alcorn State University
Marc Smith, Sinclair Community College
Frederick Spiegel, University of Arkansas
Frederick Sproull, La Roche College
Bob Sullivan, Marist College
Mark Sutherland, Hendrix College
Toure Thompson, Alabama A&M University
Scott Thomson, University of Wisconsin – Parkside
Allison van de Meene, University of Melbourne
Mary White, Southeastern Louisiana University
Steven Wilt, Bellarmine University
James Wise, Hampton University
Renna Wolfe
Virginia Young, Mercer University
Leslie Zeman, University of Washington
Daniel Zurek, Pittsburg State University
Shobhana Natarajan, Alcon Laboratories, Inc.

Introduction

class="introduction" This NASA image is a composite of several satellite-based views of Earth. To make the whole-Earth image, NASA scientists combine observations of different parts of the planet. (credit: NASA/GSFC/NOAA/USGS)

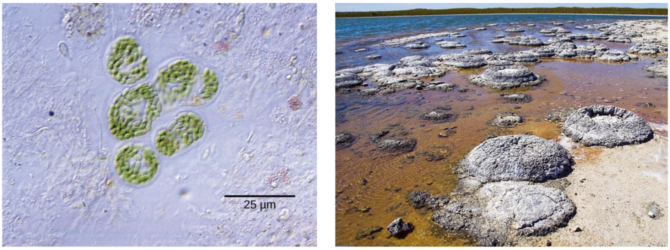
Viewed from space, Earth offers no clues about the diversity of life forms that reside there. Scientists believe that the first forms of life on Earth were microorganisms that existed for billions of years in the ocean before plants and animals appeared. The mammals, birds, and flowers so familiar to us are all relatively recent, originating 130 to 250 million years ago. The earliest representatives of the genus *Homo*, to which we belong, have inhabited this planet for only the last 2.5 million years, and only in the last 300,000 years have humans started looking like we do today.

The Science of Biology

By the end of this section, you will be able to do the following:

- Identify the shared characteristics of the natural sciences
- Summarize the steps of the scientific method
- Compare inductive reasoning with deductive reasoning
- Describe the goals of basic science and applied science

Formerly called blue-green algae, these (a) cyanobacteria, magnified 300x under a light microscope, are some of Earth's oldest life forms. These (b) stromatolites along the shores of Lake Thetis in Western Australia are ancient structures formed by layering cyanobacteria in shallow waters. (credit a: modification of work by NASA; credit b: modification of work by Ruth Ellison; scale-bar data from Matt Russell)

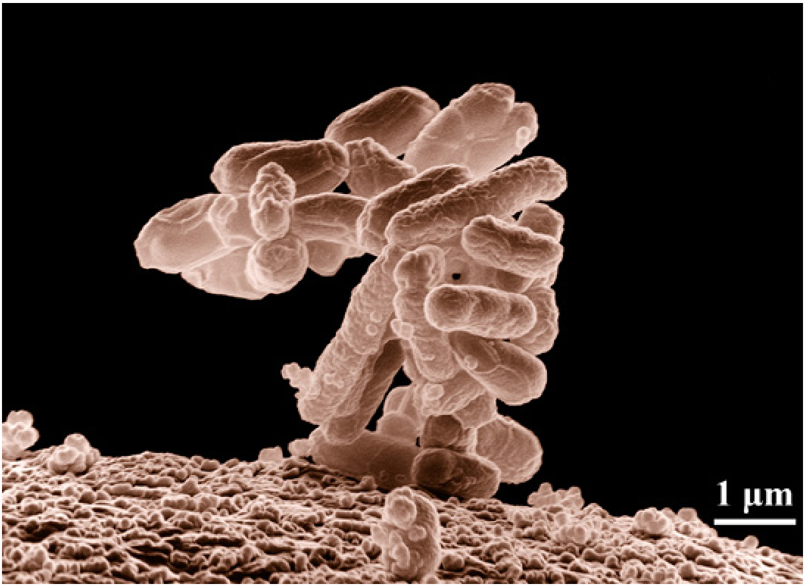


(a) (b)

What is biology? In simple terms, **biology** is the study of living organisms and their interactions with one another and their environments. This is a very

broad definition because the scope of biology is vast. Biologists may study anything from the microscopic or submicroscopic view of a cell to ecosystems and the whole living planet ([link]). Listening to the daily news, you will quickly realize how many aspects of biology we discuss every day. For example, recent news topics include *Escherichia coli* ([link]) outbreaks in spinach and *Salmonella* contamination in peanut butter. Other subjects include efforts toward finding a cure for AIDS, Alzheimer's disease, and cancer. On a global scale, many researchers are committed to finding ways to protect the planet, solve environmental issues, and reduce the effects of climate change. All of these diverse endeavors are related to different facets of the discipline of biology.

Escherichia coli (*E. coli*) bacteria, in this scanning electron micrograph, are normal residents of our digestive tracts that aid in absorbing vitamin K and other nutrients. However, virulent strains are sometimes responsible for disease outbreaks. (credit: Eric Erbe, digital colorization by Christopher Pooley, both of USDA, ARS, EMU)



The diversity of scientific fields includes astronomy, biology, computer science, geology, logic, physics, chemistry, mathematics, and many other fields. (credit: “Image Editor”/Flickr)

The Process of Science

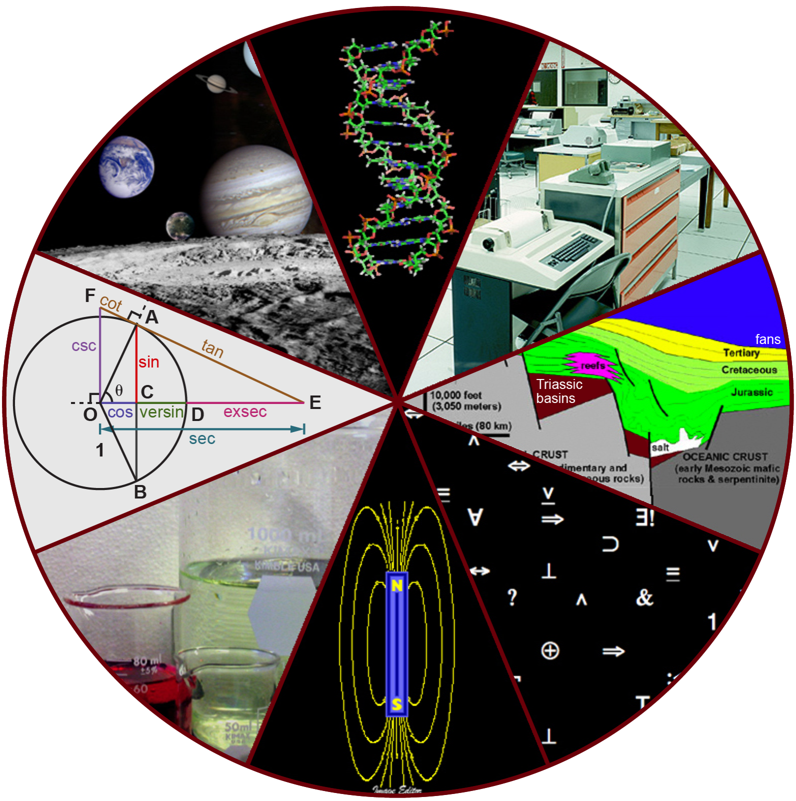
Biology is a science, but what exactly is science? What does the study of biology share with other scientific disciplines? We can define **science** (from the Latin *scientia*, meaning “knowledge”) as knowledge that covers general truths or the operation of general laws, especially when acquired and tested by the scientific method. It becomes clear from this definition that applying scientific method plays a major role in science. The **scientific method** is a method of research with defined steps that include experiments and careful observation.

We will examine scientific method steps in detail later, but one of the most important aspects of this method is the testing of hypotheses by means of repeatable experiments. A **hypothesis** is a suggested explanation for an event, which one can test. Although using the scientific method is inherent to science, it is inadequate in determining what science is. This is because it is relatively easy to apply the scientific method to disciplines such as physics and chemistry, but when it comes to disciplines like archaeology, psychology, and geology, the scientific method becomes less applicable as repeating experiments becomes more difficult.

These areas of study are still sciences, however. Consider archaeology—even though one cannot perform repeatable experiments, hypotheses may still be supported. For instance, an archaeologist can hypothesize that an ancient culture existed based on finding a piece of pottery. He or she could make further hypotheses about various characteristics of this culture, which could be correct or false through continued support or contradictions from other findings. A hypothesis may become a verified theory. A **theory** is a tested and confirmed explanation for observations or phenomena. Therefore, we may be better off to define science as fields of study that attempt to comprehend the nature of the universe.

Natural Sciences

What would you expect to see in a museum of natural sciences? Frogs? Plants? Dinosaur skeletons? Exhibits about how the brain functions? A planetarium? Gems and minerals? Maybe all of the above? Science includes such diverse fields as astronomy, biology, computer sciences, geology, logic, physics, chemistry, and mathematics ([link]). However, scientists consider those fields of science related to the physical world and its phenomena and processes **natural sciences**. Thus, a museum of natural sciences might contain any of the items listed above.



There is no complete agreement when it comes to defining what the natural sciences include, however. For some experts, the natural sciences are astronomy, biology, chemistry, earth science, and physics. Other scholars choose to divide natural sciences into **life sciences**, which study living things and include biology, and **physical sciences**, which study nonliving matter and include astronomy, geology, physics, and chemistry. Some disciplines such as biophysics and biochemistry build on both life and physical sciences and are interdisciplinary. Some refer to natural sciences as “hard science” because they rely on the use of quantitative data. Social sciences that study society and human behavior are more likely to use qualitative assessments to drive investigations and findings.

Not surprisingly, the natural science of biology has many branches or subdisciplines. Cell biologists study cell structure and function, while biologists who study anatomy investigate the structure of an entire organism. Those biologists studying physiology, however, focus on the internal functioning of an organism. Some areas of biology focus on only particular types of living things. For example, botanists explore plants, while zoologists specialize in animals.

Scientific Reasoning

One thing is common to all forms of science: an ultimate goal "to know." Curiosity and inquiry are the driving forces for the development of science. Scientists seek to understand the world and the way it operates. To do this, they use two methods of logical thinking: inductive reasoning and deductive reasoning.

Inductive reasoning is a form of logical thinking that uses related observations to arrive at a general conclusion. This type of reasoning is common in descriptive science. A life scientist such as a biologist makes observations and records them. These data can be qualitative or quantitative, and one can supplement the raw data with drawings, pictures, photos, or videos. From many observations, the scientist can infer conclusions (inductions) based on evidence. Inductive reasoning involves formulating generalizations inferred from careful observation and analyzing a large amount of data. Brain studies provide an example. In this type of research, scientists observe many live brains while people are engaged in a specific activity, such as viewing images of food. The scientist then predicts the part of the brain that "lights up" during this activity to be the part controlling the response to the selected stimulus, in this case, images of food. Excess absorption of radioactive sugar derivatives by active areas of the brain causes the various areas to "light up". Scientists use a scanner to observe the resultant increase in radioactivity. Then, researchers

can stimulate that part of the brain to see if similar responses result.

Deductive reasoning or deduction is the type of logic used in hypothesis-based science. In deductive reason, the pattern of thinking moves in the opposite direction as compared to inductive reasoning. **Deductive reasoning** is a form of logical thinking that uses a general principle or law to forecast specific results. From those general principles, a scientist can extrapolate and predict the specific results that would be valid as long as the general principles are valid. Studies in climate change can illustrate this type of reasoning. For example, scientists may predict that if the climate becomes warmer in a particular region, then the distribution of plants and animals should change.

Both types of logical thinking are related to the two main pathways of scientific study: descriptive science and hypothesis-based science. **Descriptive (or discovery) science**, which is usually inductive, aims to observe, explore, and discover, while **hypothesis-based science**, which is usually deductive, begins with a specific question or problem and a potential answer or solution that one can test. The boundary between these two forms of study is often blurred, and most scientific endeavors combine both approaches. The fuzzy boundary becomes apparent when thinking about how easily observation can lead to specific questions. For

example, a gentleman in the 1940s observed that the burr seeds that stuck to his clothes and his dog's fur had a tiny hook structure. On closer inspection, he discovered that the burrs' gripping device was more reliable than a zipper. He eventually experimented to find the best material that acted similar, and produced the hook-and-loop fastener popularly known today as Velcro. Descriptive science and hypothesis-based science are in continuous dialogue.

Historians credit Sir Francis Bacon (1561–1626) as the first to define the scientific method. (credit: Paul van Somer)

The Scientific Method

Biologists study the living world by posing questions about it and seeking science-based responses. Known as scientific method, this approach is common to other sciences as well. The scientific method was used even in ancient times, but England's Sir Francis Bacon (1561–1626) first documented it ([link]). He set up inductive methods for scientific inquiry. The scientific method is not used only by biologists; researchers from almost all fields of study can apply it as a logical, rational problem-solving method.

The scientific process typically starts with an observation (often a problem to solve) that leads to a question. Let’s think about a simple problem that starts with an observation and apply the scientific method to solve the problem. One Monday morning, a student arrives at class and quickly discovers that the classroom is too warm. That is an observation that also describes a problem: the classroom is too warm. The student then asks a question: “Why is the classroom so warm?”

Proposing a Hypothesis

Recall that a hypothesis is a suggested explanation that one can test. To solve a problem, one can propose several hypotheses. For example, one hypothesis might be, “The classroom is warm because no one turned on the air conditioning.” However, there could be other responses to the question, and therefore one may propose other hypotheses. A second hypothesis might be, “The classroom is warm because there is a power failure, and so the air conditioning doesn’t work.”

Once one has selected a hypothesis, the student can make a prediction. A prediction is similar to a hypothesis but it typically has the format “If . . . then . . . .” For example, the prediction for the first hypothesis might be, “*If* the student turns on the air conditioning, *then* the classroom will no longer be too warm.”

Testing a Hypothesis

A valid hypothesis must be testable. It should also be **falsifiable**, meaning that experimental results can disprove it. Importantly, science does not claim to “prove” anything because scientific understandings are always subject to modification with further information. This step—openness to disproving ideas—is what distinguishes sciences from non-sciences. The presence of the supernatural,

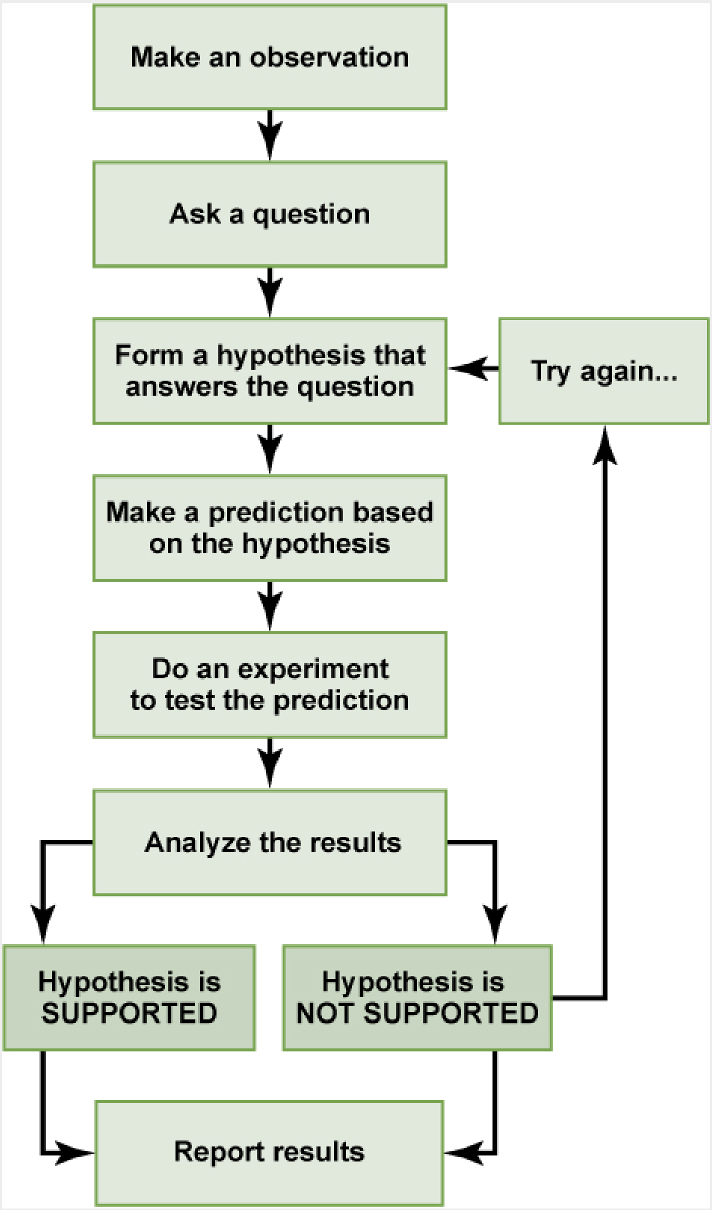
for instance, is neither testable nor falsifiable. To test a hypothesis, a researcher will conduct one or more experiments designed to eliminate one or more of the hypotheses. Each experiment will have one or more variables and one or more controls. A **variable** is any part of the experiment that can vary or change during the experiment. The **control group** contains every feature of the experimental group except it is not given the manipulation that the researcher hypothesizes. Therefore, if the experimental group's results differ from the control group, the difference must be due to the hypothesized manipulation, rather than some outside factor. Look for the variables and controls in the examples that follow. To test the first hypothesis, the student would find out if the air conditioning is on. If the air conditioning is turned on but does not work, there should be another reason, and the student should reject this hypothesis. To test the second hypothesis, the student could check if the lights in the classroom are functional. If so, there is no power failure and the student should reject this hypothesis. The students should test each hypothesis by carrying out appropriate experiments. Be aware that rejecting one hypothesis does not determine whether or not one can accept the other hypotheses. It simply eliminates one hypothesis that is not valid ([link]). Using the scientific method, the student rejects the hypotheses that are inconsistent with experimental data.

While this “warm classroom” example is based on observational results, other hypotheses and experiments might have clearer controls. For instance, a student might attend class on Monday and realize she had difficulty concentrating on the lecture. One observation to explain this occurrence might be, “When I eat breakfast before class, I am better able to pay attention.” The student could then design an experiment with a control to test this hypothesis.

In hypothesis-based science, researchers predict specific results from a general premise. We call this type of reasoning deductive reasoning: deduction proceeds from the general to the particular. However, the reverse of the process is also possible: sometimes, scientists reach a general conclusion from a number of specific observations. We call this type of reasoning inductive reasoning, and it proceeds from the particular to the general. Researchers often use inductive and deductive reasoning in tandem to advance scientific knowledge ([link]).

Visual Connection

The scientific method consists of a series of well-defined steps. If a hypothesis is not supported by experimental data, one can propose a new hypothesis.



In the example below, the scientific method is used to solve an everyday problem. Match the scientific method steps (numbered items) with the process of solving the everyday problem (lettered items).

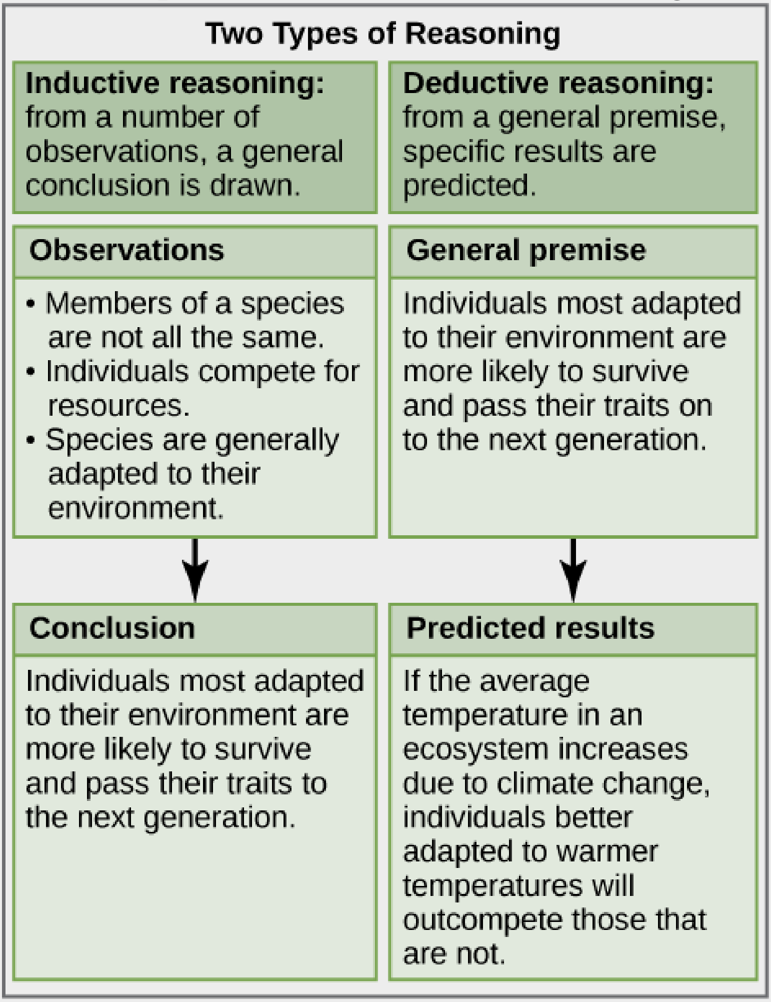
Based on the results of the experiment, is the hypothesis correct? If it is incorrect, propose some alternative hypotheses.

1. Observation	a. There is something wrong with the electrical outlet.
2. Question	b. If something is wrong with the outlet, my coffeemaker also won't work when plugged into it.
3. Hypothesis (answer)	c. My toaster doesn't toast my bread.
4. Prediction	d. I plug my coffee maker into the outlet.
5. Experiment	e. My coffeemaker works.
6. Result	f. Why doesn't my toaster work?

Visual Connection

Scientists use two types of reasoning, inductive and deductive reasoning, to advance scientific knowledge. As is the case in this example, the

conclusion from inductive reasoning can often become the premise for deductive reasoning.



Decide if each of the following is an example of inductive or deductive reasoning.

1. All flying birds and insects have wings. Birds and insects flap their wings as they move through the air. Therefore, wings enable flight.

2. Insects generally survive mild winters better than harsh ones. Therefore, insect pests will become more problematic if global temperatures increase.
3. Chromosomes, the carriers of DNA, are distributed evenly between the daughter cells during cell division. Therefore, each daughter cell will have the same chromosome set as the mother cell.
4. Animals as diverse as humans, insects, and wolves all exhibit social behavior. Therefore, social behavior must have an evolutionary advantage.

The scientific method may seem too rigid and structured. It is important to keep in mind that, although scientists often follow this sequence, there is flexibility. Sometimes an experiment leads to conclusions that favor a change in approach. Often, an experiment brings entirely new scientific questions to the puzzle. Many times, science does not operate in a linear fashion. Instead, scientists continually draw inferences and make generalizations, finding patterns as their research proceeds. Scientific reasoning is more complex than the scientific method alone suggests. Notice, too, that we can apply the scientific method to solving problems that aren't necessarily scientific in nature. After Hurricane Irma struck the Caribbean and

Florida in 2017, thousands of baby squirrels like this one were thrown from their nests. Thanks to applied science, scientists knew how to rehabilitate the squirrel. (credit: audreyjm529, Flickr) The Human Genome Project was a 13-year collaborative effort among researchers working in several different science fields. Researchers completed the project, which sequenced the entire human genome, in 2003. (credit: the U.S. Department of Energy Genome Programs (http://genomics.energy.gov)

Two Types of Science: Basic Science and Applied Science

The scientific community has been debating for the last few decades about the value of different types of science. Is it valuable to pursue science for the sake of simply gaining knowledge, or does scientific knowledge only have worth if we can apply it to solving a specific problem or to bettering our lives? This question focuses on the differences between two types of science: basic science and applied science.

Basic science or “pure” science seeks to expand knowledge regardless of the short-term application of that knowledge. It is not focused on developing a product or a service of immediate public or commercial value. The immediate goal of basic science is knowledge for knowledge’s sake, although this does not mean that, in the end, it may not result

in a practical application.

In contrast, **applied science** or “technology,” aims to use science to solve real-world problems, making it possible, for example, to improve a crop yield, find a cure for a particular disease, or save animals threatened by a natural disaster ([link]). In applied science, the problem is usually defined for the researcher.

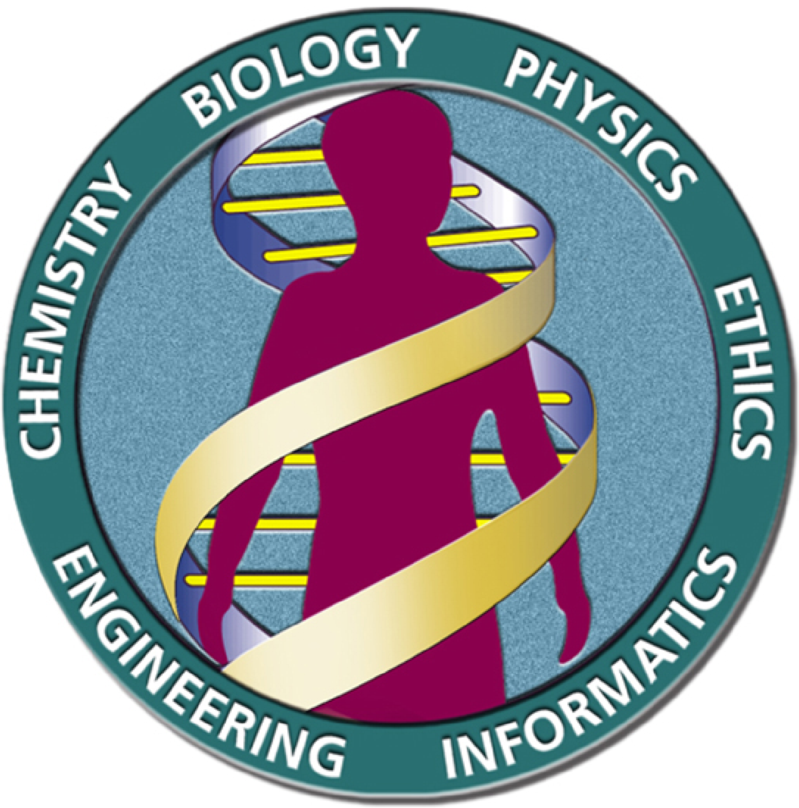
Some individuals may perceive applied science as “useful” and basic science as “useless.” A question these people might pose to a scientist advocating knowledge acquisition would be, “What for?” However, a careful look at the history of science reveals that basic knowledge has resulted in many remarkable applications of great value. Many

scientists think that a basic understanding of science is necessary before researchers develop an application therefore, applied science relies on the results that researchers generate through basic science. Other scientists think that it is time to move on from basic science in order to find solutions to actual problems. Both approaches are valid. It is true that there are problems that demand immediate attention; however, scientists would find few solutions without the help of the wide knowledge foundation that basic science generates.

One example of how basic and applied science can work together to solve practical problems occurred after the discovery of DNA structure led to an understanding of the molecular mechanisms governing DNA replication. DNA strands, unique in every human, are in our cells, where they provide the instructions necessary for life. When DNA replicates, it produces new copies of itself, shortly before a cell divides. Understanding DNA replication mechanisms enabled scientists to develop laboratory techniques that researchers now use to identify genetic diseases, pinpoint individuals who were at a crime scene, and determine paternity. Without basic science, it is unlikely that applied science would exist.

Another example of the link between basic and applied research is the Human Genome Project, a study in which researchers analyzed and mapped

each human chromosome to determine the precise sequence of DNA subunits and each gene's exact location. (The gene is the basic unit of heredity. An individual’s complete collection of genes is his or her genome.) Researchers have studied other less complex organisms as part of this project in order to gain a better understanding of human chromosomes. The Human Genome Project ([link]) relied on basic research with simple organisms and, later, with the human genome. An important end goal eventually became using the data for applied research, seeking cures and early diagnoses for genetically related diseases.



While scientists usually carefully plan research efforts in both basic science and applied science, note that some discoveries are made by **serendipity**, that is, by means of a fortunate accident or a lucky surprise. Scottish biologist Alexander Fleming discovered penicillin when he accidentally left a petri dish of *Staphylococcus* bacteria open. An unwanted mold grew on the dish, killing the bacteria. Fleming's curiosity to investigate the reason behind the bacterial death, followed by his experiments, led to the discovery of the antibiotic penicillin, which is produced by the fungus *Penicillium*. Even in the highly organized world of science, luck—when combined with an observant, curious mind—can lead to unexpected breakthroughs.

Reporting Scientific Work

Whether scientific research is basic science or applied science, scientists must share their findings in order for other researchers to expand and build upon their discoveries. Collaboration with other scientists—when planning, conducting, and analyzing results—are all important for scientific research. For this reason, important aspects of a scientist’s work are communicating with peers and disseminating results to peers. Scientists can share results by presenting them at a scientific meeting or conference, but this approach can reach only the

select few who are present. Instead, most scientists present their results in peer-reviewed manuscripts that are published in scientific journals. **Peer-reviewed manuscripts** are scientific papers that a scientist's colleagues or peers review. These colleagues are qualified individuals, often experts in the same research area, who judge whether or not the scientist's work is suitable for publication. The process of peer review helps to ensure that the research in a scientific paper or grant proposal is original, significant, logical, and thorough. Grant proposals, which are requests for research funding, are also subject to peer review. Scientists publish their work so other scientists can reproduce their experiments under similar or different conditions to expand on the findings. The experimental results must be consistent with the findings of other scientists.

A scientific paper is very different from creative writing. Although creativity is required to design experiments, there are fixed guidelines when it comes to presenting scientific results. First, scientific writing must be brief, concise, and accurate. A scientific paper needs to be succinct but detailed enough to allow peers to reproduce the experiments.

The scientific paper consists of several specific sections—introduction, materials and methods, results, and discussion. This structure is sometimes called the "IMRaD" format. There are usually

acknowledgment and reference sections as well as an **abstract** (a concise summary) at the beginning of the paper. There might be additional sections depending on the type of paper and the journal where it will be published. For example, some review papers require an outline.

The **introduction** starts with brief, but broad, background information about what is known in the field. A good introduction also gives the rationale of the work. It justifies the work carried out and also briefly mentions the end of the paper, where the researcher will present the hypothesis or research question driving the research. The introduction refers to the published scientific work of others and therefore requires citations following the style of the journal. Using the work or ideas of others without proper citation is **plagiarism**.

The **materials and methods** section includes a complete and accurate description of the substances the researchers use, and the method and techniques they use to gather data. The description should be thorough enough to allow another researcher to repeat the experiment and obtain similar results, but it does not have to be verbose. This section will also include information on how the researchers made measurements and the types of calculations and statistical analyses they used to examine raw data. Although the materials and methods section gives an accurate description of the experiments, it does

not discuss them.

Some journals require a results section followed by a discussion section, but it is more common to combine both. If the journal does not allow combining both sections, the **results** section simply narrates the findings without any further interpretation. The researchers present results with tables or graphs, but they do not present duplicate information. In the **discussion** section, the researchers will interpret the results, describe how variables may be related, and attempt to explain the observations. It is indispensable to conduct an extensive literature search to put the results in the context of previously published scientific research. Therefore, researchers include proper citations in this section as well.

Finally, the **conclusion** section summarizes the importance of the experimental findings. While the scientific paper almost certainly answers one or more scientific questions that the researchers stated, any good research should lead to more questions. Therefore, a well-done scientific paper allows the researchers and others to continue and expand on the findings.

Review articles do not follow the IMRAD format because they do not present original scientific findings, or primary literature. Instead, they summarize and comment on findings that were

published as primary literature and typically include extensive reference sections.

Section Summary

Biology is the science that studies living organisms and their interactions with one another and their environments. Science attempts to describe and understand the nature of the universe in whole or in part by rational means. Science has many fields. Those fields related to the physical world and its phenomena are natural sciences.

Science can be basic or applied. The main goal of basic science is to expand knowledge without any expectation of short-term practical application of that knowledge. The primary goal of applied research, however, is to solve practical problems.

Science uses two types of logical reasoning. Inductive reasoning uses particular results to produce general scientific principles. Deductive reasoning is a form of logical thinking that predicts results by applying general principles. The common thread throughout scientific research is using the scientific method, a step-based process that consists of making observations, defining a problem, posing hypotheses, testing these hypotheses, and drawing one or more conclusions. The testing uses proper controls. Scientists present their results in peer-

reviewed scientific papers published in scientific journals. A scientific research paper consists of several well-defined sections: introduction, materials and methods, results, and, finally, a concluding discussion. Review papers summarize the conducted research in a particular field over a period of time.

Visual Connection Questions

[link] In the example below, the scientific method is used to solve an everyday problem. Match the scientific method steps (numbered items) with the process of solving the everyday problem (lettered items). Based on the results of the experiment, is the hypothesis correct? If it is incorrect, propose some alternative hypotheses.

1. Observation	a. There is something wrong with the electrical outlet.
2. Question	b. If something is wrong with the outlet, my coffeemaker also won't work when

	plugged into it.
3. Hypothesis (answer)	c. My toaster doesn't toast my bread.
4. Prediction	d. I plug my coffee maker into the outlet.
5. Experiment	e. My coffeemaker works.
6. Result	f. Why doesn't my toaster work?

[link] 1: C; 2: F; 3: A; 4: B; 5: D; 6: E. The original hypothesis is incorrect, as the coffeemaker works when plugged into the outlet. Alternative hypotheses include that the toaster might be broken or that the toaster wasn't turned on.

[link] Decide if each of the following is an example of inductive or deductive reasoning.

1. All flying birds and insects have wings. Birds and insects flap their wings as they move through the air. Therefore, wings enable flight.
2. Insects generally survive mild winters better than harsh ones. Therefore, insect pests will become more problematic if global temperatures increase.
3. Chromosomes, the carriers of DNA,

separate into daughter cells during cell division. Therefore, each daughter cell will have the same chromosome set as the mother cell.
4. Animals as diverse as humans, insects, and wolves all exhibit social behavior. Therefore, social behavior must have an evolutionary advantage.

[link] 1: inductive; 2: deductive; 3: deductive; 4: inductive.

Review Questions

The first forms of life on Earth were _____.

1. plants
2. microorganisms
3. birds
4. dinosaurs

B

A suggested and testable explanation for an event is called a _____.

1. hypothesis
2. variable
3. theory
4. control

A

Which of the following sciences is not considered a natural science?

1. biology
2. astronomy
3. physics
4. computer science

D

The type of logical thinking that uses related observations to arrive at a general conclusion is called _____.

1. deductive reasoning
2. the scientific method
3. hypothesis-based science
4. inductive reasoning

D

The process of ______ helps to ensure that a scientist's research is original, significant, logical, and thorough.

1. publication
2. public speaking
3. peer review
4. the scientific method

C

A person notices that her houseplants that are regularly exposed to music seem to grow more quickly than those in rooms with no music. As a result, she determines that plants grow better when exposed to music. This example most closely resembles which type of reasoning?

1. inductive reasoning
2. deductive reasoning
3. neither, because no hypothesis was made
4. both inductive and deductive reasoning

A

Critical Thinking Questions

Although the scientific method is used by most of the sciences, it can also be applied to everyday situations. Think about a problem that you may have at home, at school, or with your car, and apply the scientific method to solve it.

Answers will vary, but should apply the steps of the scientific method. One possibility could be a car which doesn’t start. The hypothesis could be that the car doesn’t start because the battery is dead. The experiment would be to change the battery or to charge the battery and then check whether the car starts or not. If it starts, the problem was due to the battery, and the hypothesis is accepted.

Give an example of how applied science has had a direct effect on your daily life.

Answers will vary. One example of how applied science has had a direct effect on daily life is the presence of vaccines. Vaccines to prevent diseases such polio, measles, tetanus, and even influenza affect daily life by contributing to individual and societal health.

Name two topics that are likely to be studied by biologists, and two areas of scientific study that

would fall outside the realm of biology.

Answers will vary. Topics that fall inside the area of biological study include how diseases affect human bodies, how pollution impacts a species' habitat, and how plants respond to their environments. Topics that fall outside of biology (the "study of life") include how metamorphic rock is formed and how planetary orbits function.

Thinking about the topic of cancer, write a basic science question and an applied science question that a researcher interested in this topic might ask.

Answers will vary. Basic science: What evolutionary purpose might cancer serve? Applied science: What strategies might be found to prevent cancer from reproducing at the cellular level?

Glossary

abstract
: opening section of a scientific paper that summarizes the research and conclusions

applied science
: form of science that aims to solve real-world problems

basic science
: science that seeks to expand knowledge and understanding regardless of the short-term application of that knowledge

biology
: the study of living organisms and their interactions with one another and their environments

conclusion
: section of a scientific paper that summarizes the importance of the experimental findings

control
: part of an experiment that does not change during the experiment

deductive reasoning
: form of logical thinking that uses a general inclusive statement to forecast specific results

descriptive science
: (also, discovery science) form of science that aims to observe, explore, and investigate

discussion
: section of a scientific paper in which the

author interprets experimental results, describes how variables may be related, and attempts to explain the phenomenon in question

falsifiable
: able to be disproven by experimental results

hypothesis
: suggested explanation for an observation, which one can test

hypothesis-based science
: form of science that begins with a specific question and potential testable answers

inductive reasoning
: form of logical thinking that uses related observations to arrive at a general conclusion

introduction
: opening section of a scientific paper, which provides background information about what was known in the field prior to the research reported in the paper

life science
: field of science, such as biology, that studies living things

materials and methods
: section of a scientific paper that includes a

complete description of the substances, methods, and techniques that the researchers used to gather data

natural science
: field of science that is related to the physical world and its phenomena and processes

peer-reviewed manuscript
: scientific paper that a scientist's colleagues review who are experts in the field of study

physical science
: field of science, such as geology, astronomy, physics, and chemistry, that studies nonliving matter

plagiarism
: using other people's work or ideas without proper citation, creating the false impression that those are the author's original ideas

results
: section of a scientific paper in which the author narrates the experimental findings and presents relevant figures, pictures, diagrams, graphs, and tables, without any further interpretation

review article
: paper that summarizes and comments on findings that were published as primary

literature

science
: knowledge that covers general truths or the operation of general laws, especially when acquired and tested by the scientific method

scientific method
: method of research with defined steps that include observation, formulation of a hypothesis, testing, and confirming or falsifying the hypothesis

serendipity
: fortunate accident or a lucky surprise

theory
: tested and confirmed explanation for observations or phenomena

variable
: part of an experiment that the experimenter can vary or change

Introduction

class="introduction" Foods such as bread, fruit, and cheese are rich sources of biological macromolecules. (credit: modification of work by Bengt Nyman)

Food provides the body with the nutrients it needs to survive. Many of these critical nutrients are biological macromolecules, or large molecules, necessary for life. Different smaller organic molecule (monomer) combinations build these macromolecules (polymers). What specific biological macromolecules do living things require? How do these molecules form? What functions do they serve? We explore these questions in this chapter.

Synthesis of Biological Macromolecules
By the end of this section, you will be able to do the following:

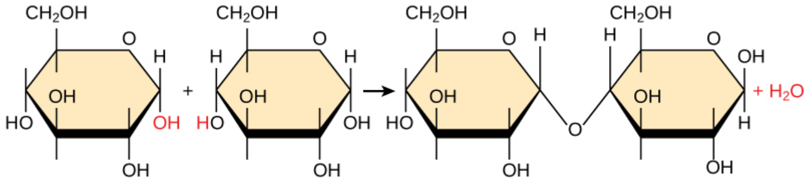
- Understand macromolecule synthesis
- Explain dehydration (or condensation) and hydrolysis reactions

As you've learned, **biological macromolecules** are large molecules, necessary for life, that are built from smaller organic molecules. There are four major biological macromolecule classes (carbohydrates, lipids, proteins, and nucleic acids). Each is an important cell component and performs a wide array of functions. Combined, these molecules make up the majority of a cell's dry mass (recall that water makes up the majority of its complete mass). Biological macromolecules are organic, meaning they contain carbon. In addition, they may contain hydrogen, oxygen, nitrogen, and additional minor elements.
In the dehydration synthesis reaction above, two glucose molecules link to form the disaccharide maltose. In the process, it forms a water molecule.

Dehydration Synthesis

Most macromolecules are made from single subunits, or building blocks, called **monomers**. The monomers combine with each other using covalent

bonds to form larger molecules known as **polymers**. In doing so, monomers release water molecules as byproducts. This type of reaction is **dehydration synthesis**, which means “to put together while losing water.”



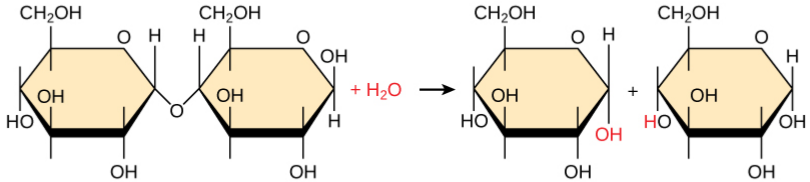
In a dehydration synthesis reaction ([link]), the hydrogen of one monomer combines with the hydroxyl group of another monomer, releasing a water molecule. At the same time, the monomers share electrons and form covalent bonds. As additional monomers join, this chain of repeating monomers forms a polymer. Different monomer types can combine in many configurations, giving rise to a diverse group of macromolecules. Even one kind of monomer can combine in a variety of ways to form several different polymers. For example, glucose monomers are the constituents of starch, glycogen, and cellulose.

In the hydrolysis reaction here, the disaccharide maltose breaks down to form two glucose monomers by adding a water molecule. Note that this reaction is the reverse of the synthesis reaction in [link].

Hydrolysis

Polymers break down into monomers during

hydrolysis. A chemical reaction occurs when inserting a water molecule across the bond. Breaking a covalent bond with this water molecule in the compound achieves this ([link]). During these reactions, the polymer breaks into two components: one part gains a hydrogen atom (H+) and the other gains a hydroxyl molecule (OH–) from a split water molecule.



Dehydration and **hydrolysis reactions** are catalyzed, or “sped up,” by specific enzymes; dehydration reactions involve the formation of new bonds, requiring energy, while hydrolysis reactions break bonds and release energy. These reactions are similar for most macromolecules, but each monomer and polymer reaction is specific for its class. For example, catalytic enzymes in the digestive system hydrolyze or break down the food we ingest into smaller molecules. This allows cells in our body to easily absorb nutrients in the intestine. A specific enzyme breaks down each macromolecule. For instance, amylase, sucrase, lactase, or maltase break down carbohydrates. Enzymes called proteases, such as pepsin and peptidase, and hydrochloric acid break down proteins. Lipases break down lipids. These broken down macromolecules provide energy for cellular activities.

Link to Learning

Visit this site to see visual representations of dehydration synthesis and hydrolysis.

Section Summary

Proteins, carbohydrates, nucleic acids, and lipids are the four major classes of biological macromolecules—large molecules necessary for life that are built from smaller organic molecules. Macromolecules are comprised of single units scientists call monomers that are joined by covalent bonds to form larger polymers. The polymer is more than the sum of its parts: it acquires new characteristics, and leads to an osmotic pressure that is much lower than that formed by its ingredients. This is an important advantage in maintaining cellular osmotic conditions. A monomer joins with another monomer with water molecule release, leading to a covalent bond forming. Scientists call these dehydration or condensation reactions. When polymers break down into smaller units (monomers), they use a water molecule for each bond broken by these reactions. Such reactions are hydrolysis reactions. Dehydration and hydrolysis reactions are similar for all macromolecules, but each monomer and polymer

reaction is specific to its class. Dehydration reactions typically require an investment of energy for new bond formation, while hydrolysis reactions typically release energy by breaking bonds.

Review Questions

Dehydration synthesis leads to formation of

1. monomers
2. polymers
3. water and polymers
4. none of the above

C

During the breakdown of polymers, which of the following reactions takes place?

1. hydrolysis
2. dehydration
3. condensation
4. covalent bond

A

The following chemical reactants produce the ester ethyl ethanoate ($C_4H_8O_2$):

C_2H_6O + CH_3COOH

What type of reaction occurs to make ethyl ethanoate?

1. condensation
2. hydrolysis
3. combustion
4. acid-base reaction

A

Critical Thinking Questions

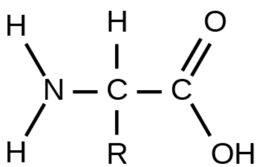
Why are biological macromolecules considered organic?

Biological macromolecules are organic because they contain carbon.

What role do electrons play in dehydration synthesis and hydrolysis?

In a dehydration synthesis reaction, the hydrogen of one monomer combines with the hydroxyl group of another monomer, releasing a molecule of water. This creates an opening in the outer shells of atoms in the monomers, which can share electrons and form covalent bonds.

Amino acids have the generic structure seen below, where R represents different carbon-based side chains.



Describe how the structure of amino acids allows them to be linked into long peptide chains to form proteins.

Amino acids can be linked into long chains through condensation reactions. One of the hydrogen atoms bonded to the nitrogen atom of an amino acid reacts with the –OH group attached to the terminal carbon on another amino acid. Since both ends of the molecule can participate in condensation reactions, peptide bonds can be made in both directions to

create a long amino acid chain.

Glossary

biological macromolecule
: large molecule necessary for life that is built from smaller organic molecules

dehydration synthesis
: (also, condensation) reaction that links monomer molecules, releasing a water molecule for each bond formed

hydrolysis
: reaction that causes breakdown of larger molecules into smaller molecules by utilizing water

monomer
: smallest unit of larger molecules that are polymers

polymer
: chain of monomer residues that covalent bonds link; polymerization is the process of polymer formation from monomers by condensation

Carbohydrates

By the end of this section, you will be able to do the following:

- Discuss the role of carbohydrates in cells and in the extracellular materials of animals and plants
- Explain carbohydrate classifications
- List common monosaccharides, disaccharides, and polysaccharides

Most people are familiar with carbohydrates, one type of macromolecule, especially when it comes to what we eat. To lose weight, some individuals adhere to “low-carb” diets. Athletes, in contrast, often “carb-load” before important competitions to ensure that they have enough energy to compete at a high level. Carbohydrates are, in fact, an essential part of our diet. Grains, fruits, and vegetables are all natural carbohydrate sources that provide energy to the body, particularly through glucose, a simple sugar that is a component of **starch** and an ingredient in many staple foods. Carbohydrates also have other important functions in humans, animals, and plants.

Scientists classify monosaccharides based on the position of their carbonyl group and the number of carbons in the backbone. Aldoses have a carbonyl group (indicated in green) at the end of the carbon chain, and ketoses have a carbonyl group in the middle of the carbon chain. Trioses, pentoses, and

hexoses have three-, five-, and six- carbon backbones, respectively. Five and six carbon monosaccharides exist in equilibrium between linear and ring forms. When the ring forms, the side chain it closes on locks into an α or β position. Fructose and ribose also form rings, although they form five-membered rings as opposed to the six-membered ring of glucose. Sucrose forms when a glucose monomer and a fructose monomer join in a dehydration reaction to form a glycosidic bond. In the process, a water molecule is lost. By convention, the carbon atoms in a monosaccharide are numbered from the terminal carbon closest to the carbonyl group. In sucrose, a glycosidic linkage forms between carbon 1 in glucose and carbon 2 in fructose. Common disaccharides include maltose (grain sugar), lactose (milk sugar), and sucrose (table sugar). Amylose and amylopectin are two different starch forms. Unbranched glucose monomer chains comprise amylose by α 1-4 glycosidic linkages. Unbranched glucose monomer chains comprise amylopectin by α 1-4 and α 1-6 glycosidic linkages. Because of the way the subunits are joined, the glucose chains have a helical structure. Glycogen (not shown) is similar in structure to amylopectin but more highly branched. In cellulose, glucose monomers are linked in unbranched chains by β 1-4 glycosidic linkages. Because of the way the glucose subunits are joined, every glucose monomer is flipped relative to the next one resulting in a linear, fibrous structure.

Insects have a hard outer exoskeleton made of chitin, a type of polysaccharide. (credit: Louise Docker)

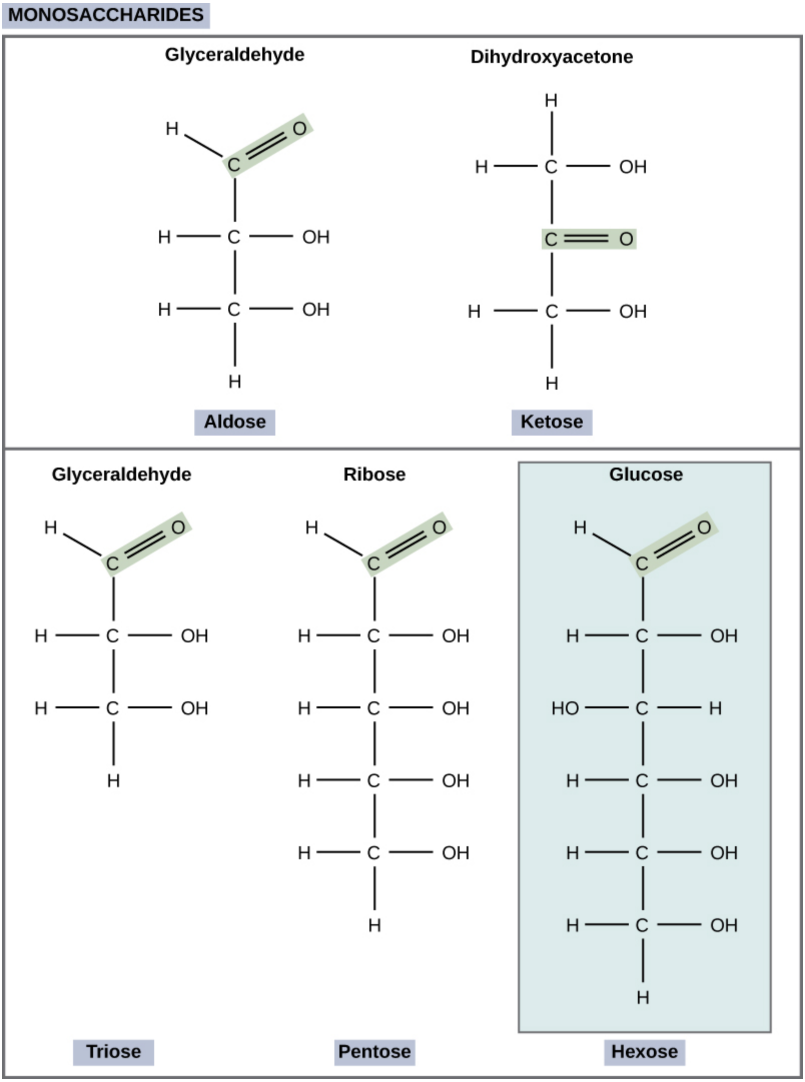
Molecular Structures

The stoichiometric formula $(CH_2O)_n$, where *n* is the number of carbons in the molecule represents **carbohydrates**. In other words, the ratio of carbon to hydrogen to oxygen is 1:2:1 in carbohydrate molecules. This formula also explains the origin of the term “carbohydrate”: the components are carbon (“carbo”) and the components of water (hence, “hydrate”). Scientists classify carbohydrates into three subtypes: monosaccharides, disaccharides, and polysaccharides.

Monosaccharides

Monosaccharides (mono- = “one”; sacchar- = “sweet”) are simple sugars, the most common of which is glucose. In monosaccharides, the number of carbons usually ranges from three to seven. Most monosaccharide names end with the suffix -ose. If the sugar has an aldehyde group (the functional group with the structure R-CHO), it is an aldose, and if it has a ketone group (the functional group with the structure RC(=O)R'), it is a ketose. Depending on the number of carbons in the sugar, they can be trioses (three carbons), pentoses (five

carbons), and/or hexoses (six carbons). [link] illustrates monosaccharides.



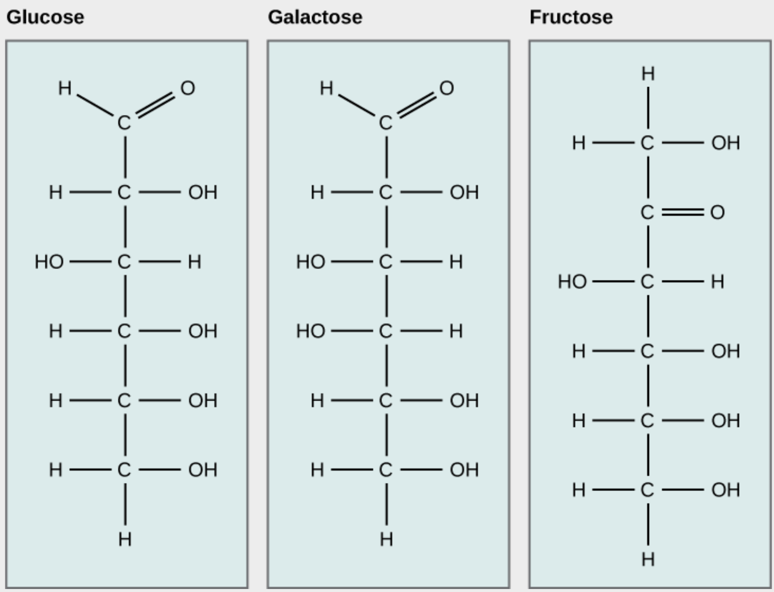
The chemical formula for glucose is $C_6H_{12}O_6$. In humans, glucose is an important source of energy. During cellular respiration, energy releases from glucose, and that energy helps make adenosine

triphosphate (ATP). Plants synthesize glucose using carbon dioxide and water, and glucose in turn provides energy requirements for the plant. Humans and other animals that feed on plants often store excess glucose as catabolized (cell breakdown of larger molecules) starch.

Galactose (part of lactose, or milk sugar) and fructose (found in sucrose, in fruit) are other common monosaccharides. Although glucose, galactose, and fructose all have the same chemical formula ($C_6H_{12}O_6$), they differ structurally and chemically (and are isomers) because of the different arrangement of functional groups around the asymmetric carbon. All these monosaccharides have more than one asymmetric carbon ([link]).

Visual Connection

Glucose, galactose, and fructose are all hexoses. They are structural isomers, meaning they have the same chemical formula ($C_6H_{12}O_6$) but a different atom arrangement.

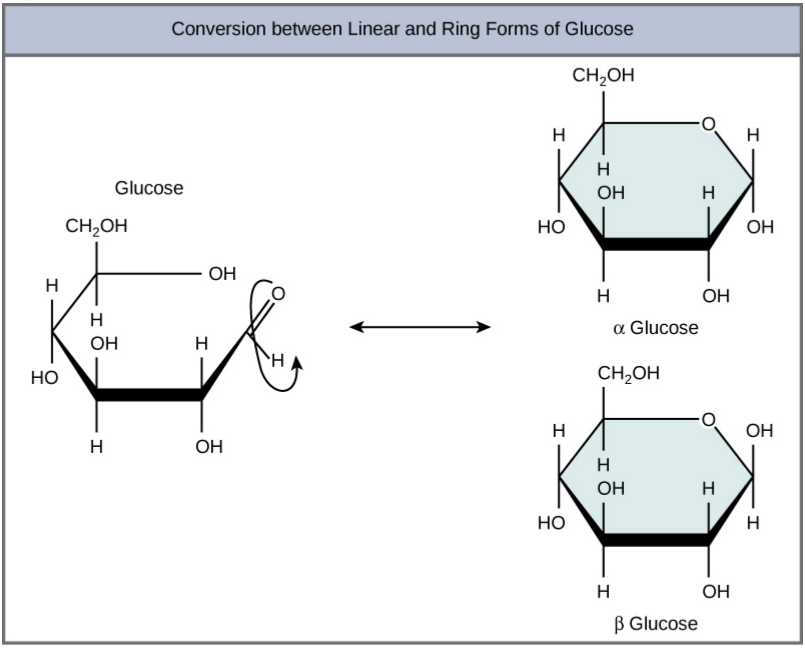


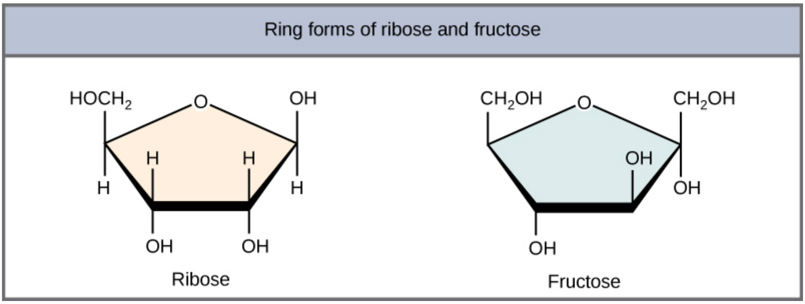
What kind of sugars are these, aldose or ketose?

Glucose, galactose, and fructose are isomeric monosaccharides (hexoses), meaning they have the same chemical formula but have slightly different structures. Glucose and galactose are aldoses, and fructose is a ketose.

Monosaccharides can exist as a linear chain or as ring-shaped molecules. In aqueous solutions they are usually in ring forms ([link]). Glucose in a ring form can have two different hydroxyl group arrangements (OH) around the anomeric carbon (carbon 1 that becomes asymmetric in the ring formation process). If the hydroxyl group is below carbon number 1 in the sugar, it is in the alpha (α)

position, and if it is above the plane, it is in the beta (β) position.

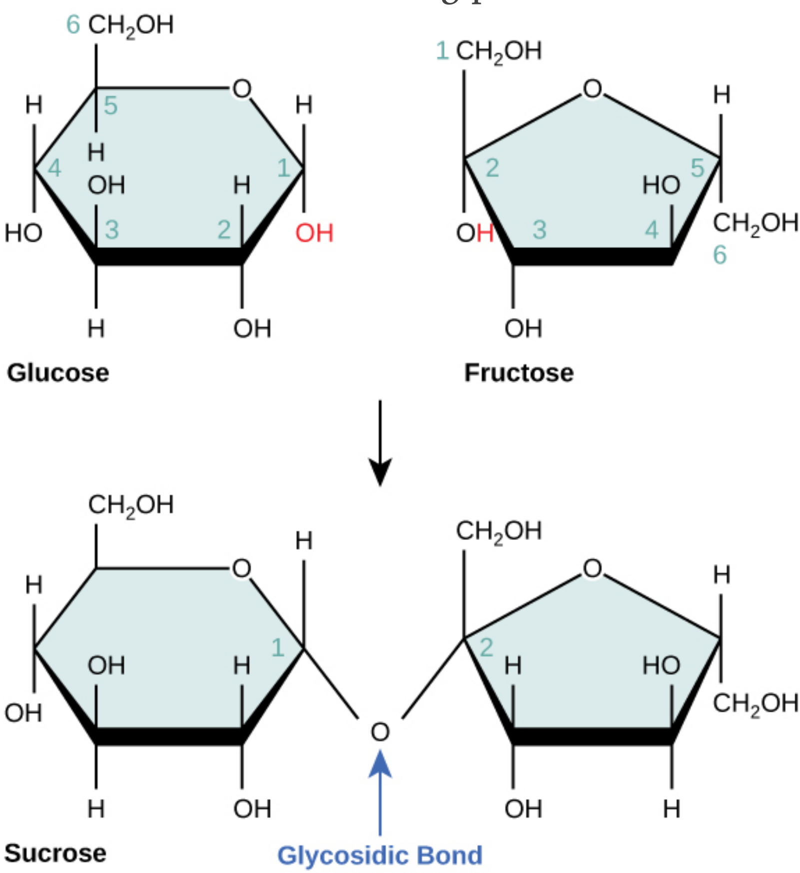




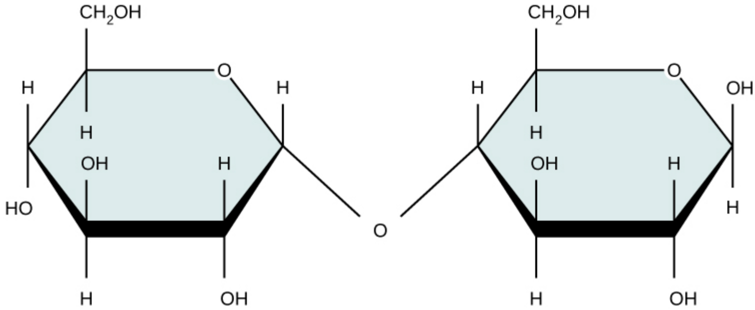
Disaccharides

Disaccharides (di- = “two”) form when two monosaccharides undergo a dehydration reaction (or a condensation reaction or dehydration synthesis). During this process, one monosaccharide's hydroxyl group combines with

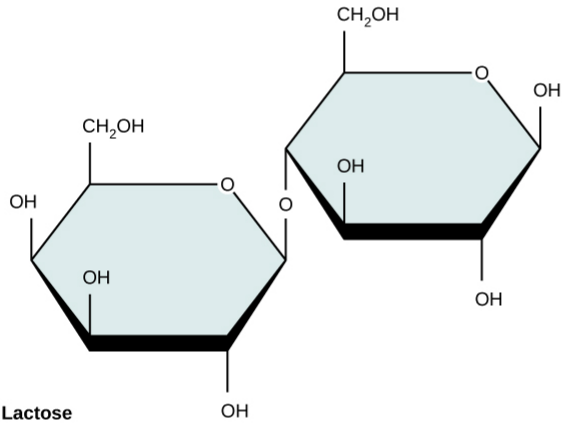
another monosaccharide's hydrogen, releasing a water molecule and forming a covalent bond. A covalent bond forms between a carbohydrate molecule and another molecule (in this case, between two monosaccharides). Scientists call this a **glycosidic bond** ([link]). Glycosidic bonds (or glycosidic linkages) can be an alpha or beta type. An alpha bond is formed when the OH group on the carbon-1 of the first glucose is below the ring plane, and a beta bond is formed when the OH group on the carbon-1 is above the ring plane.



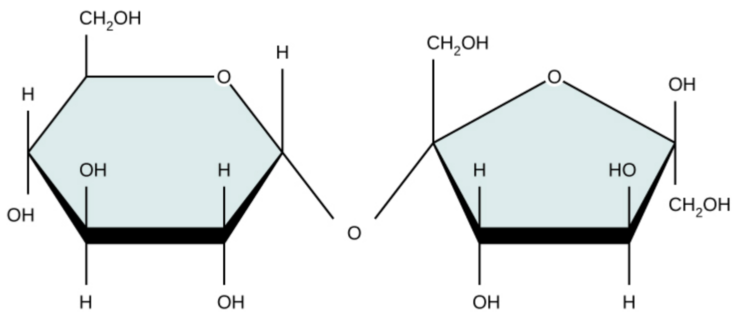
Common disaccharides include lactose, maltose, and sucrose ([link]). Lactose is a disaccharide consisting of the monomers glucose and galactose. It is naturally in milk. Maltose, or malt sugar, is a disaccharide formed by a dehydration reaction between two glucose molecules. The most common disaccharide is sucrose, or table sugar, which is comprised of glucose and fructose monomers.



Maltose



Lactose



Sucrose

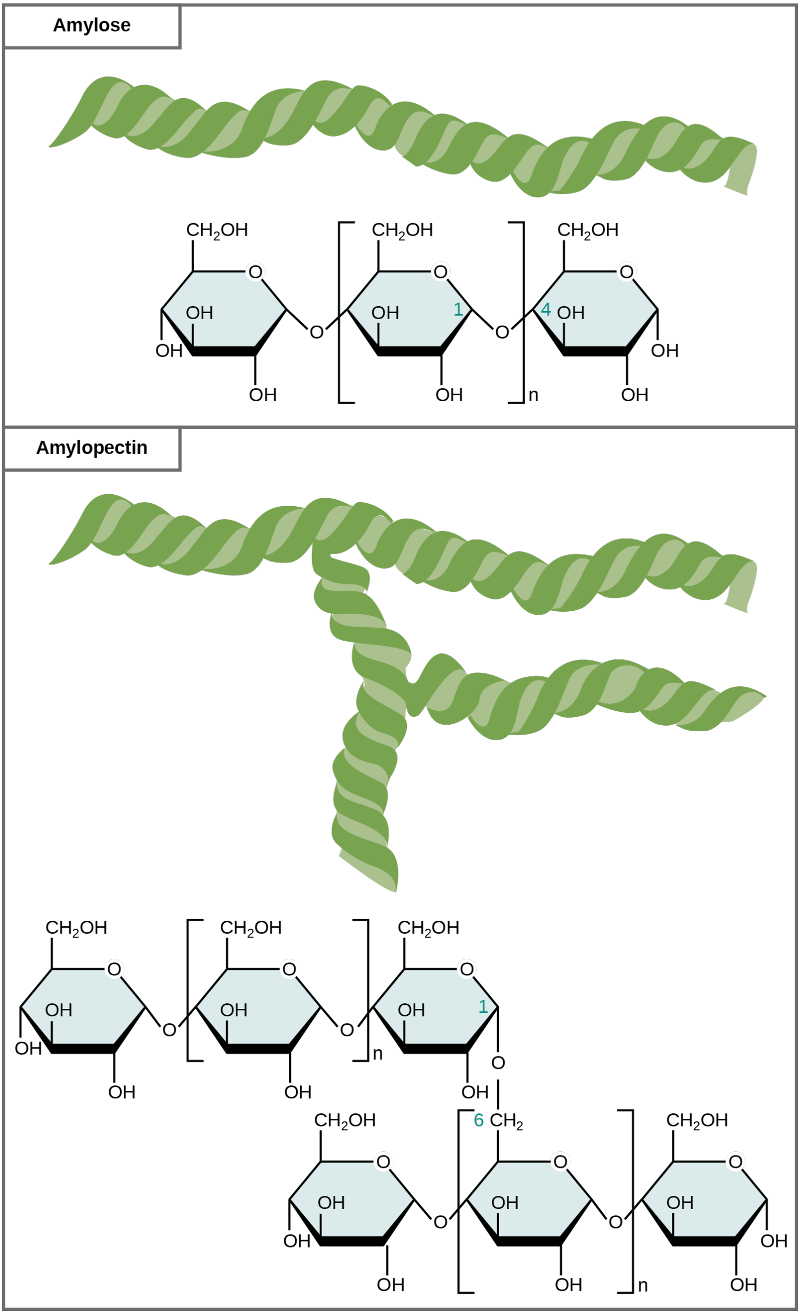
Polysaccharides

A long chain of monosaccharides linked by glycosidic bonds is a **polysaccharide** (poly- =

“many”). The chain may be branched or unbranched, and it may contain different types of monosaccharides. The molecular weight may be 100,000 daltons or more depending on the number of joined monomers. Starch, glycogen, cellulose, and chitin are primary examples of polysaccharides.

Plants store starch in the form of sugars. In plants, an amylose and amylopectic mixture (both glucose polymers) comprise these sugars. Plants are able to synthesize glucose, and they store the excess glucose, beyond their immediate energy needs, as starch in different plant parts, including roots and seeds. The starch in the seeds provides food for the embryo as it germinates and can also act as a food source for humans and animals. Enzymes break down the starch that humans consume. For example, an amylase present in saliva catalyzes, or breaks down this starch into smaller molecules, such as maltose and glucose. The cells can then absorb the glucose.

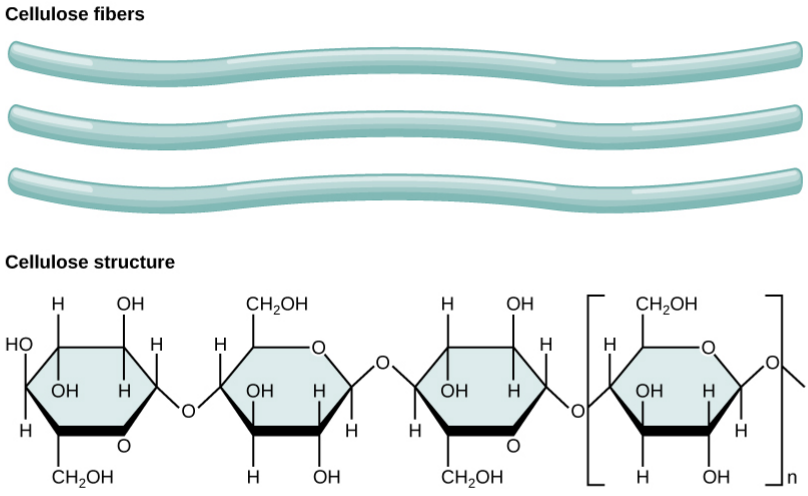
Glucose starch comprises monomers that are joined by α 1-4 or α 1-6 glycosidic bonds. The numbers 1-4 and 1-6 refer to the carbon number of the two residues that have joined to form the bond. As [link] illustrates, unbranched glucose monomer chains (only α 1-4 linkages) form the starch; whereas, amylopectin is a branched polysaccharide (α 1-6 linkages at the branch points).



Glycogen is the storage form of glucose in humans

and other vertebrates and is comprised of monomers of glucose. Glycogen is the animal equivalent of starch and is a highly branched molecule usually stored in liver and muscle cells. Whenever blood glucose levels decrease, glycogen breaks down to release glucose in a process scientists call glycogenolysis.

Cellulose is the most abundant natural biopolymer. Cellulose mostly comprises a plant's cell wall. This provides the cell structural support. Wood and paper are mostly cellulosic in nature. Glucose monomers comprise cellulose that *β* 1-4 glycosidic bonds link ([link]).



As [link] shows, every other glucose monomer in cellulose is flipped over, and the monomers are packed tightly as extended long chains. This gives cellulose its rigidity and high tensile strength—which is so important to plant cells. While human

digestive enzymes cannot break down the β 1-4 linkage, herbivores such as cows, koalas, and buffalos are able, with the help of the specialized flora in their stomach, to digest plant material that is rich in cellulose and use it as a food source. In some of these animals, certain species of bacteria and protists reside in the rumen (part of the herbivore's digestive system) and secrete the enzyme cellulase. The appendix of grazing animals also contains bacteria that digest cellulose, giving it an important role in ruminants' digestive systems. Cellulases can break down cellulose into glucose monomers that animals use as an energy source. Termites are also able to break down cellulose because of the presence of other organisms in their bodies that secrete cellulases.

Carbohydrates serve various functions in different animals. Arthropods (insects, crustaceans, and others) have an outer skeleton, the exoskeleton, which protects their internal body parts (as we see in the bee in [link]). This exoskeleton is made of the biological macromolecule **chitin**, which is a polysaccharide-containing nitrogen. It is made of repeating N-acetyl-β-d-glucosamine units, which are a modified sugar. Chitin is also a major component of fungal cell walls. Fungi are neither animals nor plants and form a kingdom of their own in the domain Eukarya.

Career Connections

Registered Dietitian

Obesity is a worldwide health concern, and many diseases such as diabetes and heart disease are becoming more prevalent because of obesity. This is one of the reasons why people increasingly seek out registered dietitians for advice. Registered dietitians help plan nutrition programs for individuals in various settings. They often work with patients in health care facilities, designing nutrition plans to treat and prevent diseases. For example, dietitians may teach a patient with diabetes how to manage blood sugar levels by

eating the correct types and amounts of carbohydrates. Dietitians may also work in nursing homes, schools, and private practices.

To become a registered dietitian, one needs to earn at least a bachelor’s degree in dietetics, nutrition, food technology, or a related field. In addition, registered dietitians must complete a supervised internship program and pass a national exam.

Those who pursue careers in dietetics take courses in nutrition, chemistry, biochemistry, biology, microbiology, and human physiology. Dietitians must become experts in the chemistry and physiology (biological functions) of food (proteins, carbohydrates, and fats).

Benefits of Carbohydrates

Are carbohydrates good for you? Some people believe that carbohydrates are bad and they should avoid them. Some diets completely forbid carbohydrate consumption, claiming that a low-carbohydrate diet helps people to lose weight faster. However, carbohydrates have been an important part of the human diet for thousands of years. Artifacts from ancient civilizations show the presence of wheat, rice, and corn in our ancestors’ storage areas.

As part of a well balanced diet, we should supplement carbohydrates with proteins, vitamins, and fats. Calorie-wise, a gram of carbohydrate provides 4.3 Kcal. For comparison, fats provide 9 Kcal/g, a less desirable ratio. Carbohydrates contain soluble and insoluble elements. The insoluble part, fiber, is mostly cellulose. Fiber has many uses. It promotes regular bowel movement by adding bulk, and it regulates the blood glucose consumption rate. Fiber also helps to remove excess cholesterol from the body. Fiber binds to the cholesterol in the small intestine, then attaches to the cholesterol and prevents the cholesterol particles from entering the bloodstream. Cholesterol then exits the body via the feces. Fiber-rich diets also have a protective role in reducing the occurrence of colon cancer. In addition, a meal containing whole grains and vegetables gives a feeling of fullness. As an immediate source of energy, glucose breaks down during the cellular respiration process, which produces ATP, the cell's energy currency. Without consuming carbohydrates, we reduce the availability of “instant energy”. Eliminating carbohydrates from the diet may be necessary for some people, but such a step may not be healthy for everyone.

Link to Learning

For an additional perspective on carbohydrates, explore “Biomolecules: the Carbohydrates” through

this interactive animation.

Section Summary

Carbohydrates are a group of macromolecules that are a vital energy source for the cell and provide structural support to plant cells, fungi, and all of the arthropods that include lobsters, crabs, shrimp, insects, and spiders. Scientists classify carbohydrates as monosaccharides, disaccharides, and polysaccharides depending on the number of monomers in the molecule. Monosaccharides are linked by glycosidic bonds that form as a result of dehydration reactions, forming disaccharides and polysaccharides with eliminating a water molecule for each bond formed. Glucose, galactose, and fructose are common monosaccharides; whereas, common disaccharides include lactose, maltose, and sucrose. Starch and glycogen, examples of polysaccharides, are the storage forms of glucose in plants and animals, respectively. The long polysaccharide chains may be branched or unbranched. Cellulose is an example of an unbranched polysaccharide; whereas, amylopectin, a constituent of starch, is a highly branched molecule. Glucose storage, in the form of polymers like starch of glycogen, makes it slightly less

accessible for metabolism; however, this prevents it from leaking out of the cell or creating a high osmotic pressure that could cause the cell to uptake excessive water.

Visual Connection Questions

[link] What kind of sugars are these, aldose or ketose?

[link] Glucose and galactose are aldoses. Fructose is a ketose.

Review Questions

An example of a monosaccharide is ______.

1. fructose
2. glucose
3. galactose
4. all of the above

D

Cellulose and starch are examples of:

1. monosaccharides
2. disaccharides
3. lipids
4. polysaccharides

D

Plant cell walls contain which of the following in abundance?

1. starch
2. cellulose
3. glycogen
4. lactose

B

Lactose is a disaccharide formed by the formation of a ______ bond between glucose and ______.

1. glycosidic; lactose
2. glycosidic; galactose
3. hydrogen; sucrose
4. hydrogen; fructose

B

Which of the following is not an extracellular matrix role of carbohydrates?

1. protect an insect’s internal organs from external trauma
2. prevent plant cells from lysing after the plant is watered
3. maintain the shape of a fungal spore
4. provide energy for muscle movement

D

Critical Thinking Questions

Describe the similarities and differences between glycogen and starch.

Glycogen and starch are polysaccharides. They are the storage form of glucose. Glycogen is stored in animals in the liver and in muscle cells, whereas starch is stored in the roots, seeds, and leaves of plants. Starch has two different forms, one unbranched (amylose) and

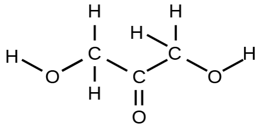
one branched (amylopectin), whereas glycogen is a single type of a highly branched molecule.

Why is it impossible for humans to digest food that contains cellulose?

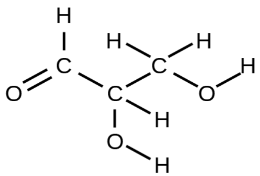
The β 1-4 glycosidic linkage in cellulose cannot be broken down by human digestive enzymes. Herbivores such as cows, koalas, and buffalos are able to digest grass that is rich in cellulose and use it as a food source because bacteria and protists in their digestive systems, especially in the rumen, secrete the enzyme cellulase. Cellulases can break down cellulose into glucose monomers that can be used as an energy source by the animal.

Draw the ketose and aldose forms of a monosaccharide with the chemical formula $C_3H_6O_3$. How is the structure of the monosaccharide changed from one form to the other in the human body?

The human body switches carbohydrates between their aldose and ketose forms using a family of enzymes called isomerases. The ketose triose is called dihydroxyacetone, and has the oxygen double-bonded to the center carbon:



The aldose is called glyceraldehyde, and can have the oxygen double-bonded to the first or third carbon of the molecule:



Glossary

carbohydrate
: biological macromolecule in which the ratio of carbon to hydrogen and to oxygen is 1:2:1; carbohydrates serve as energy sources and structural support in cells and form arthropods' cellular exoskeleton

cellulose
: polysaccharide that comprises the plants' cell wall; provides structural support to the cell

chitin

type of carbohydrate that forms the outer skeleton of all arthropods that include crustaceans and insects; it also forms fungi cell walls

disaccharide
: two sugar monomers that a glycosidic bond links

glycogen
: storage carbohydrate in animals

glycosidic bond
: bond formed by a dehydration reaction between two monosaccharides with eliminating a water molecule

monosaccharide
: single unit or monomer of carbohydrates

polysaccharide
: long chain of monosaccharides; may be branched or unbranched

starch
: storage carbohydrate in plants

Lipids

By the end of this section, you will be able to do the following:

- Describe the four major types of lipids
- Explain the role of fats in storing energy
- Differentiate between saturated and unsaturated fatty acids
- Describe phospholipids and their role in cells
- Define the basic structure of a steroid and some steroid functions
- Explain how cholesterol helps maintain the plasma membrane's fluid nature

Lipids include a diverse group of compounds that are largely nonpolar in nature. This is because they are hydrocarbons that include mostly nonpolar carbon–carbon or carbon–hydrogen bonds. Non-polar molecules are hydrophobic ("water fearing"), or insoluble in water. Lipids perform many different functions in a cell. Cells store energy for long-term use in the form of fats. Lipids also provide insulation from the environment for plants and animals ([link]). For example, they help keep aquatic birds and mammals dry when forming a protective layer over fur or feathers because of their water-repellant hydrophobic nature. Lipids are also the building blocks of many hormones and are an important constituent of all cellular membranes. Lipids include fats, oils, waxes, phospholipids, and steroids. Hydrophobic lipids in aquatic mammals' fur, such as

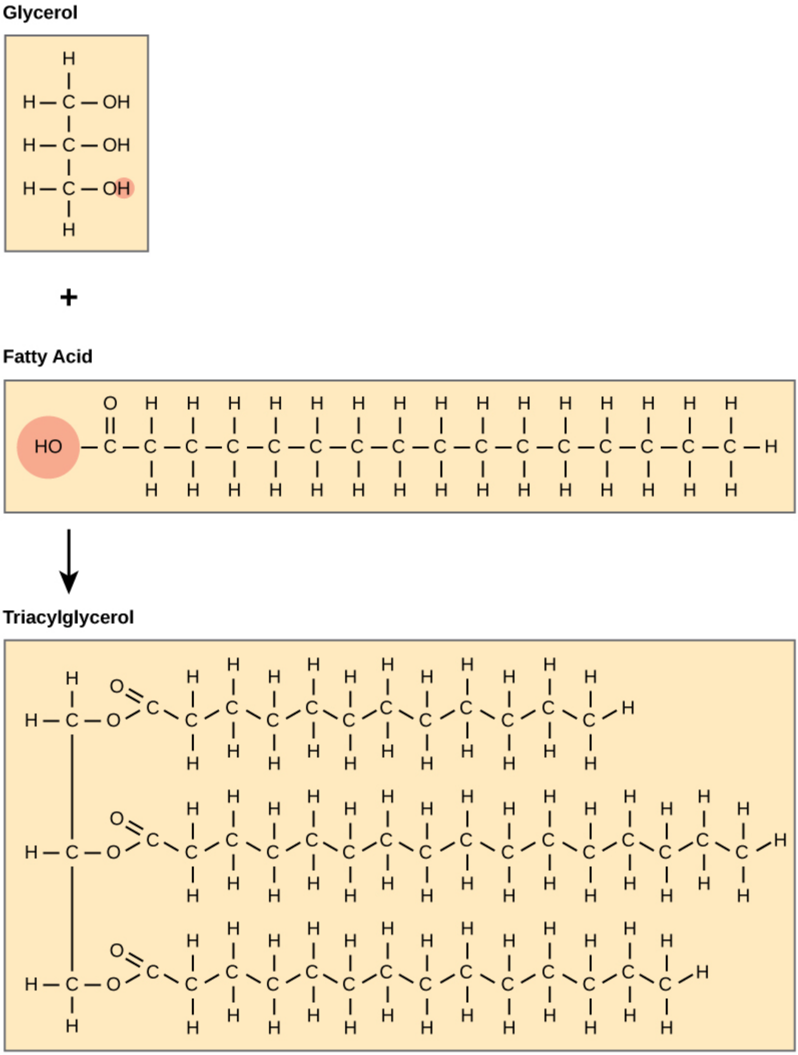
this river otter, protect them from the elements. (credit: Ken Bosma)

Joining three fatty acids to a glycerol backbone in a dehydration reaction forms triacylglycerol. Three water molecules release in the process. Stearic acid is a common saturated fatty acid. Oleic acid is a common unsaturated fatty acid. Saturated fatty acids have hydrocarbon chains connected by single bonds only. Unsaturated fatty acids have one or more double bonds. Each double bond may be in a *cis* or *trans* configuration. In the *cis* configuration, both hydrogens are on the same side of the hydrocarbon chain. In the *trans* configuration, the hydrogens are on opposite sides. A *cis* double bond causes a kink in the chain. Alpha-linolenic acid is an

example of an omega-3 fatty acid. It has three *cis* double bonds and, as a result, a curved shape. For clarity, the diagram does not show the carbons. Each singly bonded carbon has two hydrogens associated with it, which the diagram also does not show.

Fats and Oils

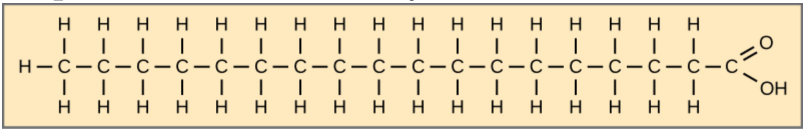
A fat molecule consists of two main components—glycerol and fatty acids. Glycerol is an organic compound (alcohol) with three carbons, five hydrogens, and three hydroxyl (OH) groups. Fatty acids have a long chain of hydrocarbons to which a carboxyl group is attached, hence the name “fatty acid.” The number of carbons in the fatty acid may range from 4 to 36. The most common are those containing 12–18 carbons. In a fat molecule, the fatty acids attach to each of the glycerol molecule's three carbons with an ester bond through an oxygen atom ([link]).



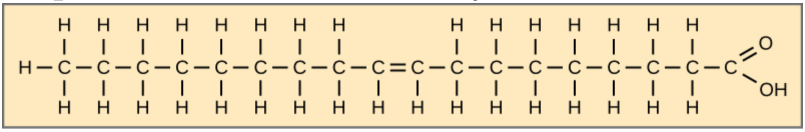
During this ester bond formation, three water molecules are released. The three fatty acids in the triacylglycerol may be similar or dissimilar. We also call fats **triacylglycerols** or **triglycerides** because of their chemical structure. Some fatty acids have common names that specify their origin. For example, palmitic acid, a **saturated fatty acid**, is

derived from the palm tree. Arachidic acid is derived from *Arachis hypogea,* the scientific name for groundnuts or peanuts.

Fatty acids may be saturated or unsaturated. In a fatty acid chain, if there are only single bonds between neighboring carbons in the hydrocarbon chain, the fatty acid is saturated. Saturated fatty acids are saturated with hydrogen. In other words, the number of hydrogen atoms attached to the carbon skeleton is maximized. Stearic acid is an example of a saturated fatty acid ([link]).



When the hydrocarbon chain contains a double bond, the fatty acid is **unsaturated**. Oleic acid is an example of an unsaturated fatty acid ([link]).



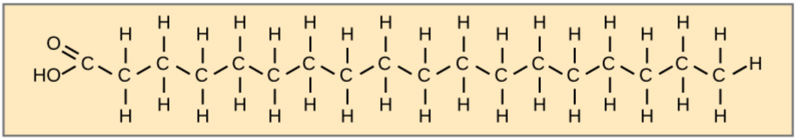
Most unsaturated fats are liquid at room temperature. We call these oils. If there is one double bond in the molecule, then it is a monounsaturated fat (e.g., olive oil), and if there is more than one double bond, then it is a polyunsaturated fat (e.g., canola oil).

When a fatty acid has no double bonds, it is a saturated fatty acid because it is not possible to add

more hydrogen to the chain's carbon atoms. A fat may contain similar or different fatty acids attached to glycerol. Long straight fatty acids with single bonds generally pack tightly and are solid at room temperature. Animal fats with stearic acid and palmitic acid (common in meat) and the fat with butyric acid (common in butter) are examples of saturated fats. Mammals store fats in specialized cells, or adipocytes, where fat globules occupy most of the cell’s volume. Plants store fat or oil in many seeds and use them as a source of energy during seedling development. Unsaturated fats or oils are usually of plant origin and contain *cis* unsaturated fatty acids. *Cis* and *trans* indicate the configuration of the molecule around the double bond. If hydrogens are present in the same plane, it is a cis fat. If the hydrogen atoms are on two different planes, it is a **trans fat**. The *cis* double bond causes a bend or a “kink” that prevents the fatty acids from packing tightly, keeping them liquid at room temperature ([link]). Olive oil, corn oil, canola oil, and cod liver oil are examples of unsaturated fats. Unsaturated fats help to lower blood cholesterol levels; whereas, saturated fats contribute to plaque formation in the arteries.

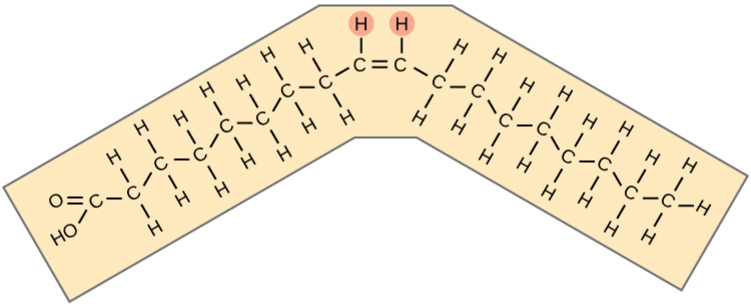
Saturated fatty acid

Stearic acid

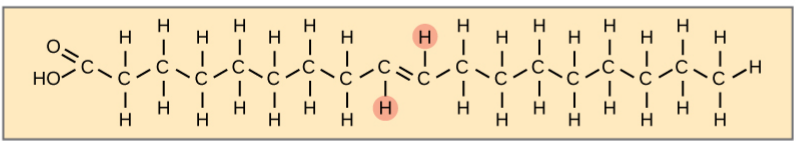


Unsaturated fatty acids

Cis oleic acid



Trans oleic acid



Trans Fats

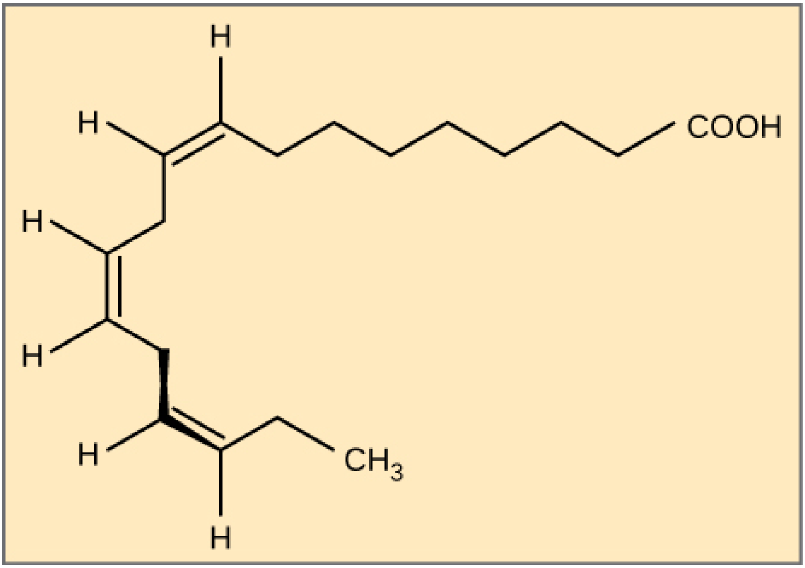
The food industry artificially hydrogenates oils to make them semi-solid and of a consistency desirable for many processed food products. Simply speaking, hydrogen gas is bubbled through oils to solidify them. During this hydrogenation process, double bonds of the *cis*- conformation in the hydrocarbon chain may convert to double bonds in the *trans*- conformation.

Margarine, some types of peanut butter, and

shortening are examples of artificially hydrogenated trans fats. Recent studies have shown that an increase in trans fats in the human diet may lead to higher levels of low-density lipoproteins (LDL), or "bad" cholesterol, which in turn may lead to plaque deposition in the arteries, resulting in heart disease. Many fast food restaurants have recently banned using trans fats, and food labels are required to display the trans fat content.

Omega Fatty Acids

Essential fatty acids are those that the human body requires but does not synthesize. Consequently, they have to be supplemented through ingestion via the diet. **Omega**-3 fatty acids (like those in [link]) fall into this category and are one of only two known for humans (the other is omega-6 fatty acid). These are polyunsaturated fatty acids and are omega-3 because a double bond connects the third carbon from the hydrocarbon chain's end to its neighboring carbon.



The farthest carbon away from the carboxyl group is numbered as the omega (ω) carbon, and if the double bond is between the third and fourth carbon from that end, it is an omega-3 fatty acid. Nutritionally important because the body does not make them, omega-3 fatty acids include alpha-linoleic acid (ALA), eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA), all of which are polyunsaturated. Salmon, trout, and tuna are good sources of omega-3 fatty acids. Research indicates that omega-3 fatty acids reduce the risk of sudden death from heart attacks, lower triglycerides in the blood, decrease blood pressure, and prevent thrombosis by inhibiting blood clotting. They also reduce inflammation, and may help lower the risk of some cancers in animals.

Like carbohydrates, fats have received considerable

bad publicity. It is true that eating an excess of fried foods and other “fatty” foods leads to weight gain. However, fats do have important functions. Many vitamins are fat soluble, and fats serve as a long-term storage form of fatty acids: a source of energy. They also provide insulation for the body. Therefore, we should consume “healthy” fats in moderate amounts on a regular basis.

Lipids comprise waxy coverings on some leaves. (credit: Roger Griffith)

Waxes

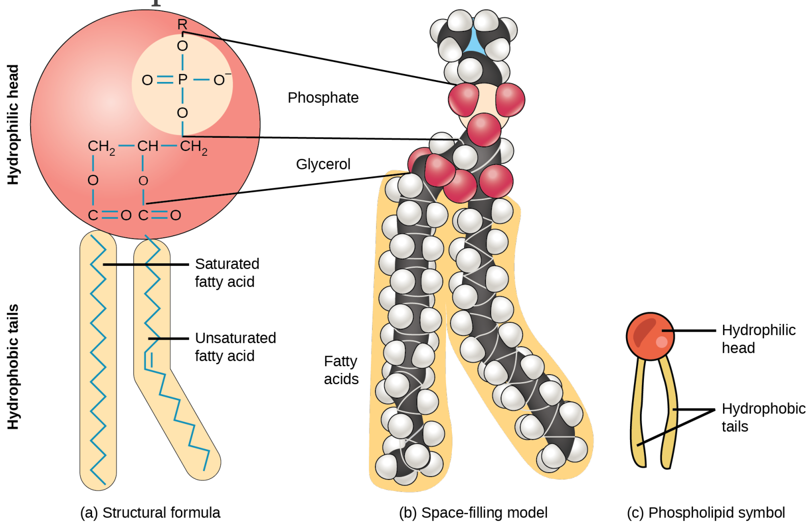
Wax covers some aquatic birds' feathers and some plants' leaf surfaces. Because of waxes' hydrophobic nature, they prevent water from sticking on the surface ([link]). Long fatty acid chains esterified to long-chain alcohols comprise waxes.

A phospholipid is a molecule with two fatty acids and a modified phosphate group attached to a glycerol backbone. Adding a charged or polar chemical group may modify the phosphate. The phospholipid bilayer is the major component of all cellular membranes. The hydrophilic head groups of the phospholipids face the aqueous solution. The hydrophobic tails are sequestered in the middle of the bilayer.

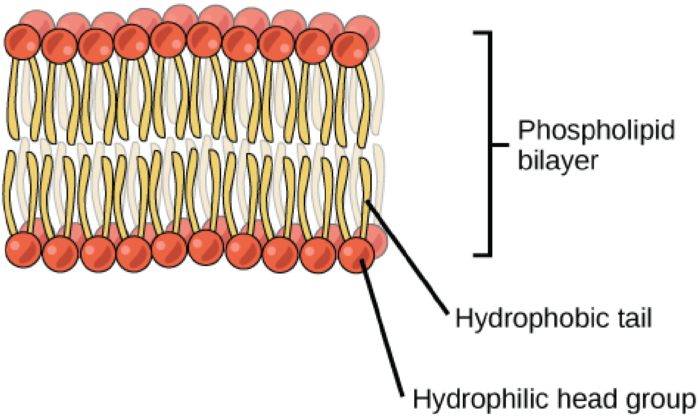
Phospholipids

Phospholipids are major plasma membrane constituents that comprise cells' outermost layer. Like fats, they are comprised of fatty acid chains attached to a glycerol or sphingosine backbone. However, instead of three fatty acids attached as in

triglycerides, there are two fatty acids forming diacylglycerol, and a modified phosphate group occupies the glycerol backbone's third carbon ([link]). A phosphate group alone attached to a diaglycerol does not qualify as a phospholipid. It is phosphatidate (diacylglycerol 3-phosphate), the precursor of phospholipids. An alcohol modifies the phosphate group. Phosphatidylcholine and phosphatidylserine are two important phospholipids that are in plasma membranes.



A phospholipid is an amphipathic molecule, meaning it has a hydrophobic and a hydrophilic part. The fatty acid chains are hydrophobic and cannot interact with water; whereas, the phosphate-containing group is hydrophilic and interacts with water ([link]).



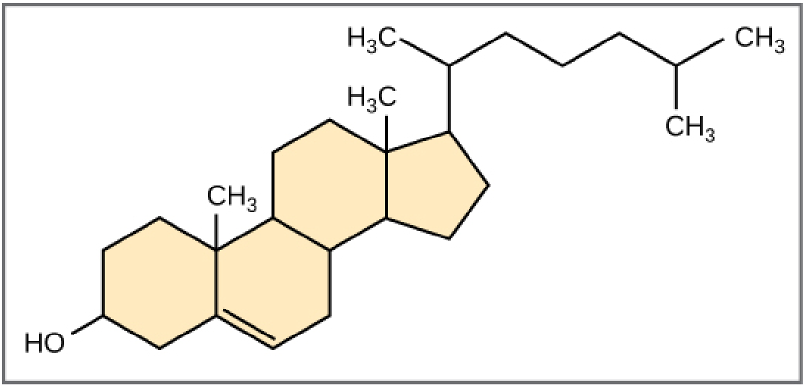
The head is the hydrophilic part, and the tail contains the hydrophobic fatty acids. In a membrane, a bilayer of phospholipids forms the structure's matrix, phospholipids' fatty acid tails face inside, away from water; whereas, the phosphate group faces the outside, aqueous side ([link]).

Phospholipids are responsible for the plasma membrane's dynamic nature. If a drop of phospholipids is placed in water, it spontaneously forms a structure that scientists call a micelle, where the hydrophilic phosphate heads face the outside and the fatty acids face the structure's interior. Four fused hydrocarbon rings comprise steroids such as cholesterol and cortisol.

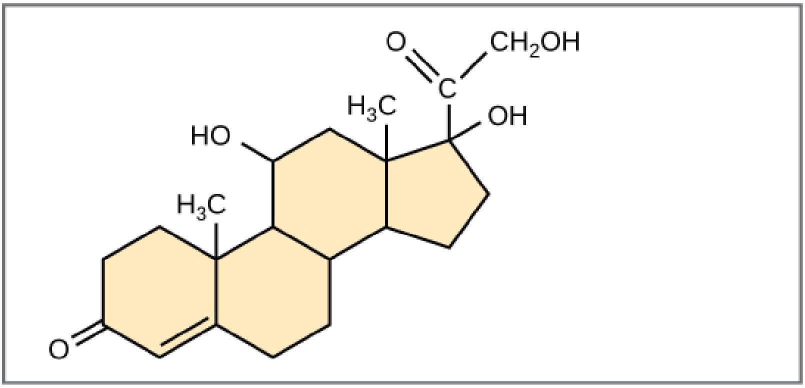
Steroids

Unlike the phospholipids and fats that we discussed earlier, **steroids** have a fused ring structure.

Although they do not resemble the other lipids, scientists group them with them because they are also hydrophobic and insoluble in water. All steroids have four linked carbon rings and several of them, like cholesterol, have a short tail ([link]). Many steroids also have the –OH functional group, which puts them in the alcohol classification (sterols).



Cholesterol



Cortisol

Cholesterol is the most common steroid. The liver synthesizes cholesterol and is the precursor to many steroid hormones such as testosterone and estradiol,

which gonads and endocrine glands secrete. It is also the precursor to Vitamin D. Cholesterol is also the precursor of bile salts, which help emulsifying fats and their subsequent absorption by cells. Although lay people often speak negatively about cholesterol, it is necessary for the body's proper functioning. Sterols (cholesterol in animal cells, phytosterol in plants) are components of the plasma membrane of cells and are found within the phospholipid bilayer.

Link to Learning

For an additional perspective on lipids, explore the interactive animation "Biomolecules: The Lipids".

Section Summary

Lipids are a class of macromolecules that are nonpolar and hydrophobic in nature. Major types include fats and oils, waxes, phospholipids, and steroids. Fats are a stored form of energy and are also known as triacylglycerols or triglycerides. Fats are comprised of fatty acids and either glycerol or sphingosine. Fatty acids may be unsaturated or saturated, depending on the presence or absence of

double bonds in the hydrocarbon chain. If only single bonds are present, they are saturated fatty acids. Unsaturated fatty acids may have one or more double bonds in the hydrocarbon chain. Phospholipids comprise the membrane's matrix. They have a glycerol or sphingosine backbone to which two fatty acid chains and a phosphate-containing group are attached. Steroids are another class of lipids. Their basic structure has four fused carbon rings. Cholesterol is a type of steroid and is an important constituent of the plasma membrane, where it helps to maintain the membrane's fluid nature. It is also the precursor of steroid hormones such as testosterone.

Review Questions

Saturated fats have all of the following characteristics except:

1. they are solid at room temperature
2. they have single bonds within the carbon chain
3. they are usually obtained from animal sources
4. they tend to dissolve in water easily

D

Phospholipids are important components of _____.

1. the plasma membrane of cells
2. the ring structure of steroids
3. the waxy covering on leaves
4. the double bond in hydrocarbon chains

A

Cholesterol is an integral part of plasma membranes. Based on its structure, where is it found in the membrane?

1. on the extracellular surface
2. embedded with the phospholipid heads
3. within the tail bilayer
4. attached to the intracellular surface

C

Critical Thinking Questions

Explain at least three functions that lipids serve

in plants and/or animals.

Fat serves as a valuable way for animals to store energy. It can also provide insulation. Waxes can protect plant leaves and mammalian fur from getting wet. Phospholipids and steroids are important components of animal cell membranes, as well as plant, fungal, and bacterial membranes.

Why have trans fats been banned from some restaurants? How are they created?

Trans fats are created artificially when hydrogen gas is bubbled through oils to solidify them. The double bonds of the *cis* conformation in the hydrocarbon chain may be converted to double bonds in the *trans* configuration. Some restaurants are banning trans fats because they cause higher levels of LDL, or “bad”cholesterol.

Why are fatty acids better than glycogen for storing large amounts of chemical energy?

Fats have a higher energy density than carbohydrates (averaging 9kcal/gram versus 4.3kcal/gram respectively). Thus, on a per gram

basis, more energy can be stored in fats than can be stored in carbohydrates. Additionally, fats are packaged into spherical globules to minimize interactions with the water-based plasma membrane, while glycogen is a large branched carbohydrate that cannot be compacted for storage.

Part of cortisol's role in the body involves passing through the plasma membrane to initiate signaling inside a cell. Describe how the structures of cortisol and the plasma membrane allow this to occur.

Cortisol is a small, generally hydrophobic molecule, while the phospholipids that create plasma membranes have a hydrophilic head and hydrophobic tails. Since cortisol is hydrophobic, it can interact with the sequestered tails of the phospholipids in the center of the plasma membrane. This, along with its small size, allows cortisol to move through the plasma membrane to the inside of the cell.

Glossary

lipid
: macromolecule that is nonpolar and insoluble

in water

omega fat
type of polyunsaturated fat that the body requires; numbering the carbon omega starts from the methyl end or the end that is farthest from the carboxylic end

phospholipid
membranes' major constituent; comprised of two fatty acids and a phosphate-containing group attached to a glycerol backbone

saturated fatty acid
long-chain hydrocarbon with single covalent bonds in the carbon chain; the number of hydrogen atoms attached to the carbon skeleton is maximized

steroid
type of lipid comprised of four fused hydrocarbon rings forming a planar structure

trans fat
fat formed artificially by hydrogenating oils, leading to a different arrangement of double bond(s) than those in naturally occurring lipids

triacylglycerol (also, triglyceride)
fat molecule; consists of three fatty acids linked to a glycerol molecule

unsaturated fatty acid
: long-chain hydrocarbon that has one or more double bonds in the hydrocarbon chain

wax
: lipid comprised of a long-chain fatty acid that is esterified to a long-chain alcohol; serves as a protective coating on some feathers, aquatic mammal fur, and leaves

Proteins
By the end of this section, you will be able to do the following:

- Describe the functions proteins perform in the cell and in tissues
- Discuss the relationship between amino acids and proteins
- Explain the four levels of protein organization
- Describe the ways in which protein shape and function are linked

Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. Proteins may be structural, regulatory, contractile, or protective. They may serve in transport, storage, or membranes; or they may be toxins or enzymes. Each cell in a living system may contain thousands of proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, amino acid polymers arranged in a linear sequence.

Types and Functions of Proteins

Enzymes, which living cells produce, are catalysts in biochemical reactions (like digestion) and are usually complex or conjugated proteins. Each

enzyme is specific for the substrate (a reactant that binds to an enzyme) upon which it acts. The enzyme may help in breakdown, rearrangement, or synthesis reactions. We call enzymes that break down their substrates catabolic enzymes. Those that build more complex molecules from their substrates are anabolic enzymes, and enzymes that affect the rate of reaction are catalytic enzymes. Note that all enzymes increase the reaction rate and, therefore, are organic catalysts. An example of an enzyme is salivary amylase, which hydrolyzes its substrate amylose, a component of starch.

Hormones are chemical-signaling molecules, usually small proteins or steroids, secreted by endocrine cells that act to control or regulate specific physiological processes, including growth, development, metabolism, and reproduction. For example, insulin is a protein hormone that helps regulate the blood glucose level. [link] lists the primary types and functions of proteins.

Protein Types and Functions

Type	Examples	Functions
Digestive Enzymes	Amylase, lipase, pepsin, trypsin	Help in food by catabolizing

		nutrients into monomeric units
Transport	Hemoglobin, albumin	Carry substances in the blood or lymph throughout the body
Structural	Actin, tubulin, keratin	Construct different structures, like the cytoskeleton
Hormones	Insulin, thyroxine	Coordinate different body systems' activity
Defense	Immunoglobulins	Protect the body from foreign pathogens
Contractile	Actin, myosin	Effect muscle contraction
Storage	Legume storage proteins, egg white (albumin)	Provide nourishment in early embryo development and the seedling

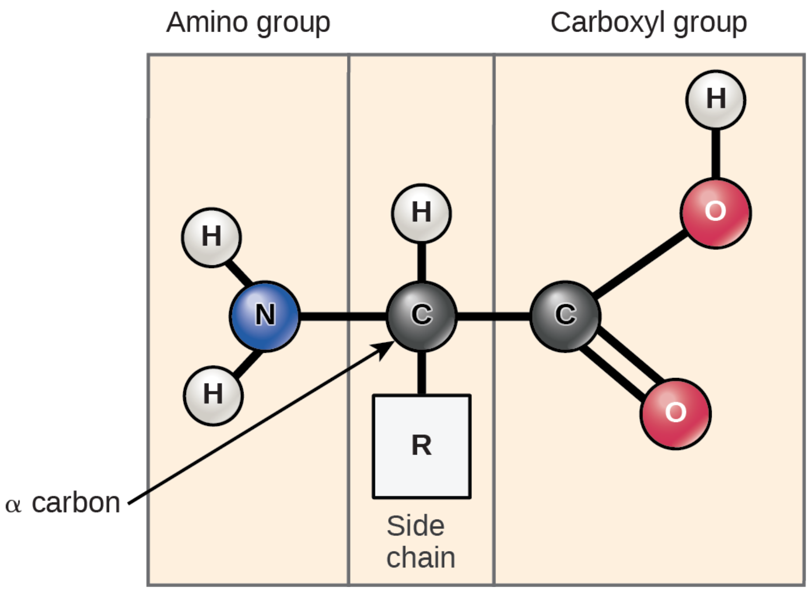
Proteins have different shapes and molecular weights. Some proteins are globular in shape; whereas, others are fibrous in nature. For example, hemoglobin is a globular protein, but collagen, located in our skin, is a fibrous protein. Protein shape is critical to its function, and many different

types of chemical bonds maintain this shape. Changes in temperature, pH, and exposure to chemicals may lead to permanent changes in the protein's shape, leading to loss of function, or **denaturation**. Different arrangements of the same 20 types of amino acids comprise all proteins. Two rare new amino acids were discovered recently (selenocystein and pirrolysine), and additional new discoveries may be added to the list.

Amino acids have a central asymmetric carbon to which an amino group, a carboxyl group, a hydrogen atom, and a side chain (R group) are attached. Peptide bond formation is a dehydration synthesis reaction. The carboxyl group of one amino acid is linked to the incoming amino acid's amino group. In the process, it releases a water molecule.

Amino Acids

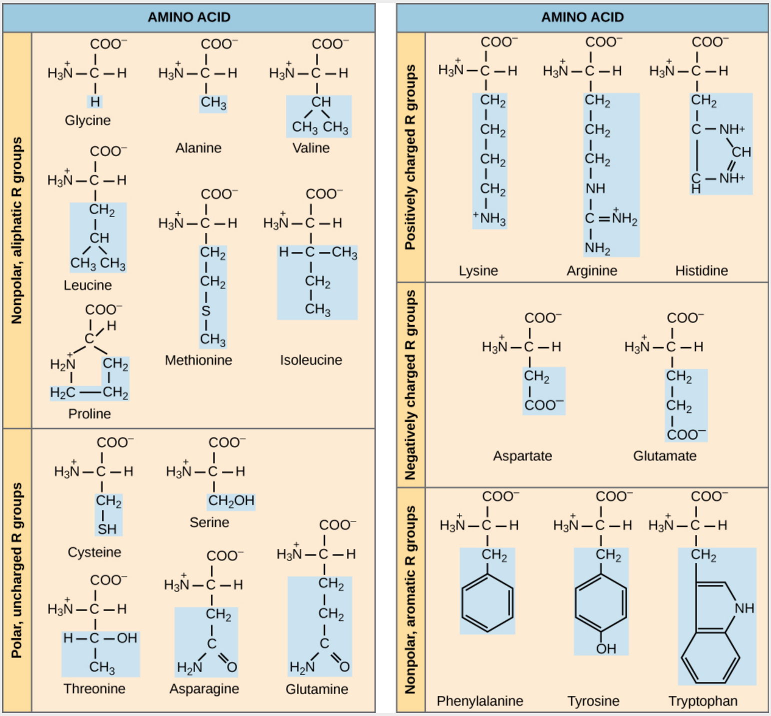
Amino acids are the monomers that comprise proteins. Each amino acid has the same fundamental structure, which consists of a central carbon atom, or the alpha (α) carbon, bonded to an amino group (NH_2), a carboxyl group (COOH), and to a hydrogen atom. Every amino acid also has another atom or group of atoms bonded to the central atom known as the R group ([link]).



Scientists use the name "amino acid" because these acids contain both amino group and carboxyl-acid-group in their basic structure. As we mentioned, there are 20 common amino acids present in proteins. Nine of these are essential amino acids in humans because the human body cannot produce them and we obtain them from our diet. For each amino acid, the R group (or side chain) is different ([link]).

Visual Connection

There are 20 common amino acids commonly found in proteins, each with a different R group (variant group) that determines its chemical nature.



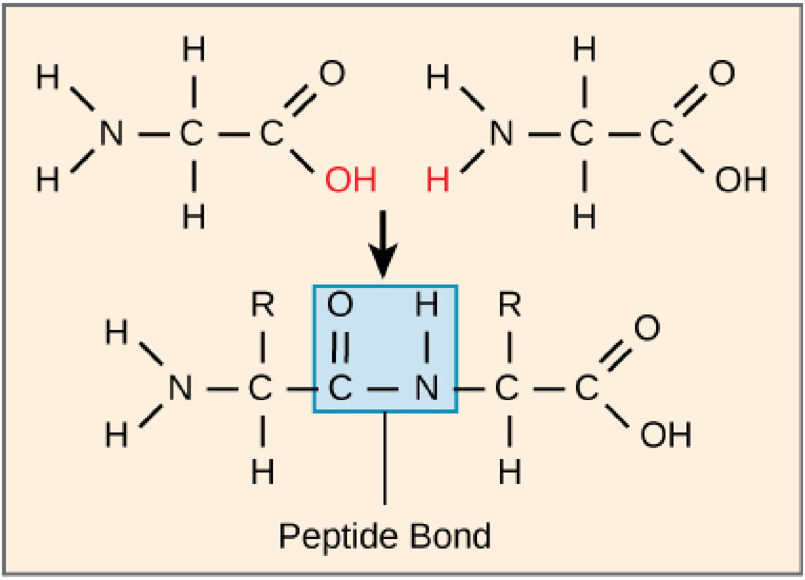
Which categories of amino acid would you expect to find on a soluble protein's surface and which would you expect to find in the interior? What distribution of amino acids would you expect to find in a protein embedded in a lipid bilayer?

The chemical nature of the side chain determines the amino acid's nature (that is, whether it is acidic, basic, polar, or nonpolar). For example, the amino acid glycine has a hydrogen atom as the R group. Amino acids such as valine, methionine, and alanine are nonpolar or hydrophobic in nature, while amino acids such as serine, threonine, and cysteine are

polar and have hydrophilic side chains. The side chains of lysine and arginine are positively charged, and therefore these amino acids are also basic amino acids. Proline has an R group that is linked to the amino group, forming a ring-like structure. Proline is an exception to the amino acid's standard structure since its amino group is not separate from the side chain ([link]).

A single upper case letter or a three-letter abbreviation represents amino acids. For example, the letter V or the three-letter symbol val represent valine. Just as some fatty acids are essential to a diet, some amino acids also are necessary. These essential amino acids in humans include isoleucine, leucine, and cysteine. Essential amino acids refer to those necessary to build proteins in the body, but not those that the body produces. Which amino acids are essential varies from organism to organism.

The sequence and the number of amino acids ultimately determine the protein's shape, size, and function. A covalent bond, or **peptide bond**, attaches to each amino acid, which a dehydration reaction forms. One amino acid's carboxyl group and the incoming amino acid's amino group combine, releasing a water molecule. The resulting bond is the peptide bond ([link]).



The products that such linkages form are peptides. As more amino acids join to this growing chain, the resulting chain is a polypeptide. Each polypeptide has a free amino group at one end. This end the N terminal, or the amino terminal, and the other end has a free carboxyl group, also the C or carboxyl terminal. While the terms polypeptide and protein are sometimes used interchangeably, a polypeptide is technically a polymer of amino acids, whereas the term protein is used for a polypeptide or polypeptides that have combined together, often have bound non-peptide prosthetic groups, have a distinct shape, and have a unique function. After protein synthesis (translation), most proteins are modified. These are known as post-translational modifications. They may undergo cleavage, phosphorylation, or may require adding other chemical groups. Only after these modifications is

the protein completely functional.

Link to Learning

Click through the steps of protein synthesis in this interactive tutorial.

Evolution Connection

The Evolutionary Significance of Cytochrome c

Cytochrome c is an important component of the electron transport chain, a part of cellular respiration, and it is normally located in the cellular organelle, the mitochondrion. This protein has a heme prosthetic group, and the heme's central ion alternately reduces and oxidizes during electron transfer. Because this essential protein's role in producing cellular energy is crucial, it has changed very little over millions of years. Protein sequencing has shown that there is a considerable amount of cytochrome c amino acid sequence homology among different species. In other words, we can assess evolutionary kinship by measuring the similarities or differences among various species' DNA or protein sequences.

Scientists have determined that human cytochrome c contains 104 amino acids. For each cytochrome c molecule from different organisms that scientists have sequenced to date, 37 of these amino acids

appear in the same position in all cytochrome c samples. This indicates that there may have been a common ancestor. On comparing the human and chimpanzee protein sequences, scientists did not find a sequence difference. When researchers compared human and rhesus monkey sequences, the single difference was in one amino acid. In another comparison, human to yeast sequencing shows a difference in the 44th position.

Bovine serum insulin is a protein hormone comprised of two peptide chains, A (21 amino acids long) and B (30 amino acids long). In each chain, three-letter abbreviations that represent the amino acids' names in the order they are present indicate primary structure. The amino acid cysteine (cys) has a sulfhydryl (SH) group as a side chain. Two sulfhydryl groups can react in the presence of oxygen to form a disulfide (S-S) bond. Two disulfide bonds connect the A and B chains together, and a third helps the A chain fold into the correct shape. Note that all disulfide bonds are the same length, but we have drawn them different sizes for clarity. The beta chain of hemoglobin is 147 residues in length, yet a single amino acid substitution leads to sickle cell anemia. In normal hemoglobin, the amino acid at position seven is glutamate. In sickle cell hemoglobin, a valine replaces glutamate. In this blood smear, visualized at 535x magnification using

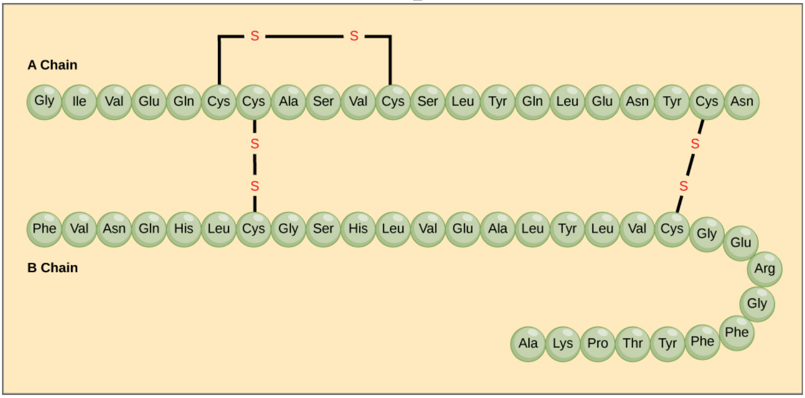
bright field microscopy, sickle cells are crescent shaped, while normal cells are disc-shaped. (credit: modification of work by Ed Uthman; scale-bar data from Matt Russell) The α-helix and β-pleated sheet are secondary structures of proteins that form because of hydrogen bonding between carbonyl and amino groups in the peptide backbone. Certain amino acids have a propensity to form an α-helix, while others have a propensity to form a β-pleated sheet. A variety of chemical interactions determine the proteins' tertiary structure. These include hydrophobic interactions, ionic bonding, hydrogen bonding, and disulfide linkages. Observe the four levels of protein structure in these illustrations. (credit: modification of work by National Human Genome Research Institute)

Protein Structure

As we discussed earlier, a protein's shape is critical to its function. For example, an enzyme can bind to a specific substrate at an active site. If this active site is altered because of local changes or changes in overall protein structure, the enzyme may be unable to bind to the substrate. To understand how the protein gets its final shape or conformation, we need to understand the four levels of protein structure: primary, secondary, tertiary, and quaternary.

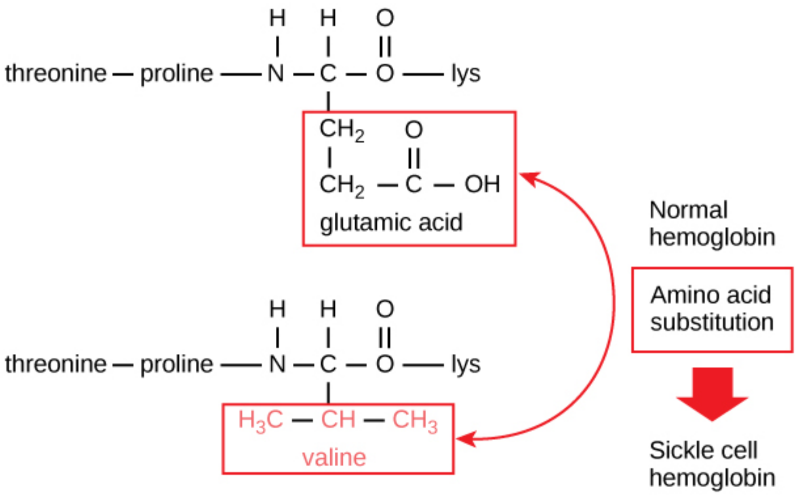
Primary Structure

Amino acids' unique sequence in a polypeptide chain is its **primary structure**. For example, the pancreatic hormone insulin has two polypeptide chains, A and B, and they are linked together by disulfide bonds. The N terminal amino acid of the A chain is glycine; whereas, the C terminal amino acid is asparagine ([link]). The amino acid sequences in the A and B chains are unique to insulin.

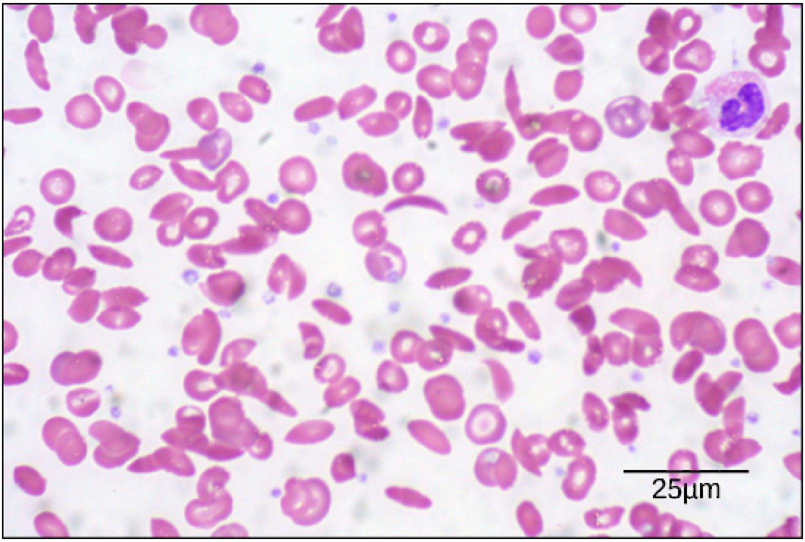


The gene encoding the protein ultimately determines the unique sequence for every protein. A change in nucleotide sequence of the gene's coding region may lead to adding a different amino acid to the growing polypeptide chain, causing a change in protein structure and function. In sickle cell anemia, the hemoglobin β chain (a small portion of which we show in [link]) has a single amino acid substitution, causing a change in protein structure and function. Specifically, valine in the β chain substitutes the amino acid glutamic. What is most remarkable to consider is that a hemoglobin molecule is comprised of two alpha and two beta chains that each consist of about 150 amino acids.

The molecule, therefore, has about 600 amino acids. The structural difference between a normal hemoglobin molecule and a sickle cell molecule—which dramatically decreases life expectancy—is a single amino acid of the 600. What is even more remarkable is that three nucleotides each encode those 600 amino acids, and a single base change (point mutation), 1 in 1800 bases causes the mutation.

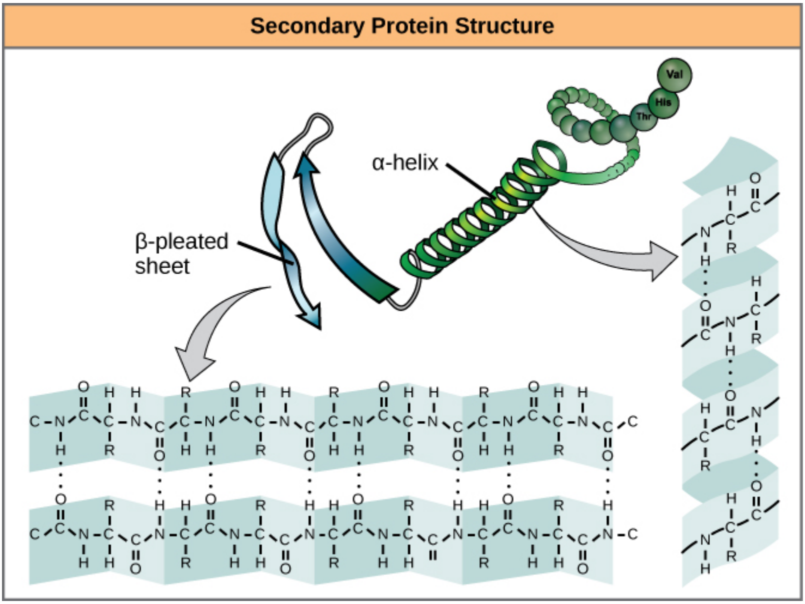


Because of this change of one amino acid in the chain, hemoglobin molecules form long fibers that distort the biconcave, or disc-shaped, red blood cells and causes them to assume a crescent or "sickle" shape, which clogs blood vessels ([link]). This can lead to myriad serious health problems such as breathlessness, dizziness, headaches, and abdominal pain for those affected by this disease.



Secondary Structure

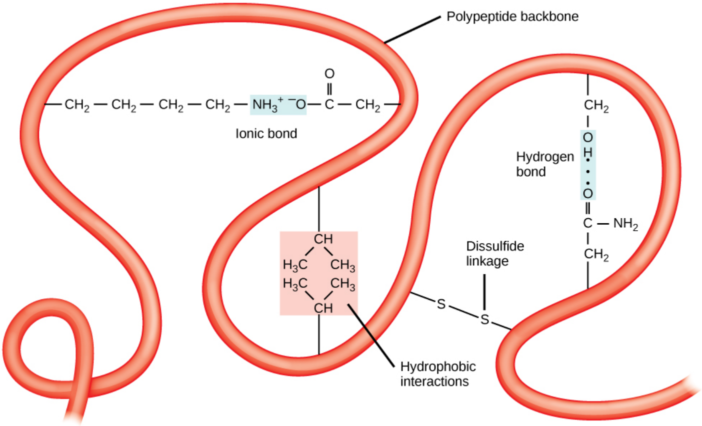
The local folding of the polypeptide in some regions gives rise to the **secondary structure** of the protein. The most common are the ***α*-helix** and ***β*-pleated sheet** structures ([link]). Both structures are held in shape by hydrogen bonds. The hydrogen bonds form between the oxygen atom in the carbonyl group in one amino acid and another amino acid that is four amino acids farther along the chain.



Every helical turn in an alpha helix has 3.6 amino acid residues. The polypeptide's R groups (the variant groups) protrude out from the α-helix chain. In the β-pleated sheet, hydrogen bonding between atoms on the polypeptide chain's backbone form the "pleats". The R groups are attached to the carbons and extend above and below the pleat's folds. The pleated segments align parallel or antiparallel to each other, and hydrogen bonds form between the partially positive nitrogen atom in the amino group and the partially negative oxygen atom in the peptide backbone's carbonyl group. The α-helix and β-pleated sheet structures are in most globular and fibrous proteins and they play an important structural role.

Tertiary Structure

The polypeptide's unique three-dimensional structure is its **tertiary structure** ([link]). This structure is in part due to chemical interactions at work on the polypeptide chain. Primarily, the interactions among R groups create the protein's complex three-dimensional tertiary structure. The nature of the R groups in the amino acids involved can counteract forming the hydrogen bonds we described for standard secondary structures. For example, R groups with like charges repel each other and those with unlike charges are attracted to each other (ionic bonds). When protein folding takes place, the nonpolar amino acids' hydrophobic R groups lie in the protein's interior; whereas, the hydrophilic R groups lie on the outside. Scientists also call the former interaction types hydrophobic interactions. Interaction between cysteine side chains forms disulfide linkages in the presence of oxygen, the only covalent bond that forms during protein folding.

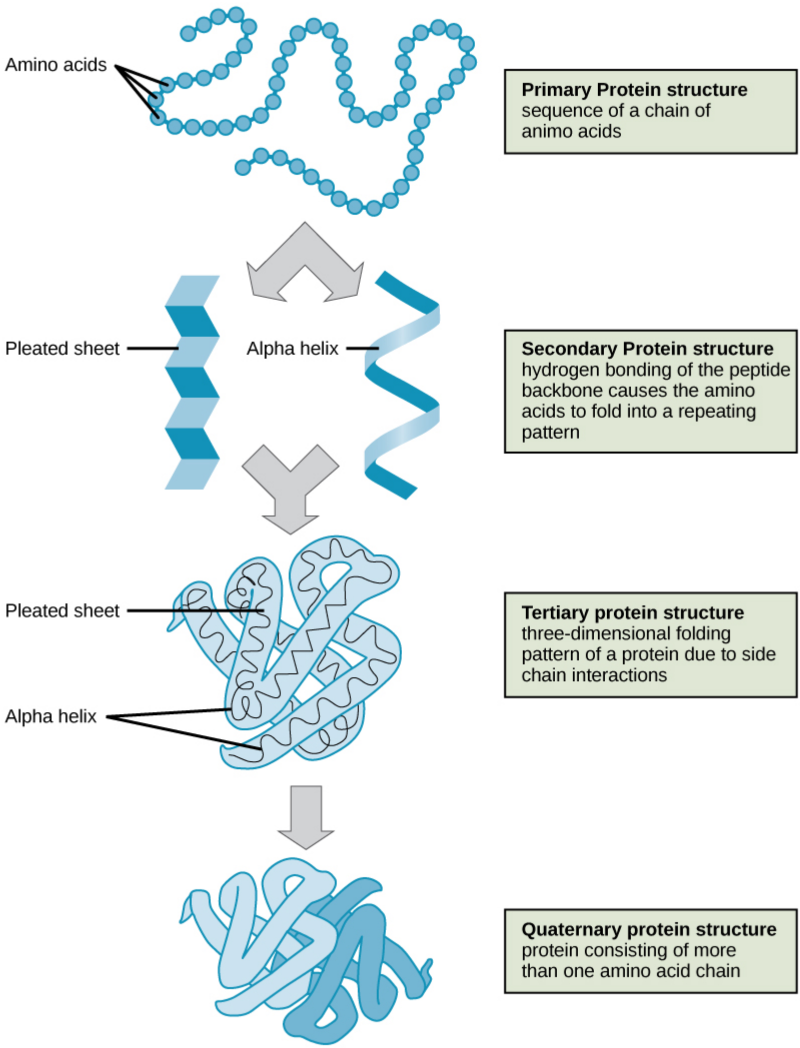


All of these interactions, weak and strong, determine the protein's final three-dimensional shape. When a protein loses its three-dimensional shape, it may no longer be functional.

Quaternary Structure

In nature, some proteins form from several polypeptides, or subunits, and the interaction of these subunits forms the **quaternary structure**. Weak interactions between the subunits help to stabilize the overall structure. For example, insulin (a globular protein) has a combination of hydrogen and disulfide bonds that cause it to mostly clump into a ball shape. Insulin starts out as a single polypeptide and loses some internal sequences in the presence of post-translational modification after forming the disulfide linkages that hold the remaining chains together. Silk (a fibrous protein), however, has a β-pleated sheet structure that is the result of hydrogen bonding between different chains.

[link] illustrates the four levels of protein structure (primary, secondary, tertiary, and quaternary).



Denaturation and Protein Folding

Each protein has its own unique sequence and shape that chemical interactions hold together. If the protein is subject to changes in temperature, pH, or

exposure to chemicals, the protein structure may change, losing its shape without losing its primary sequence in what scientists call denaturation. Denaturation is often reversible because the polypeptide's primary structure is conserved in the process if the denaturing agent is removed, allowing the protein to resume its function. Sometimes denaturation is irreversible, leading to loss of function. One example of irreversible protein denaturation is frying an egg. The albumin protein in the liquid egg white denatures when placed in a hot pan. Not all proteins denature at high temperatures. For instance, bacteria that survive in hot springs have proteins that function at temperatures close to boiling. The stomach is also very acidic, has a low pH, and denatures proteins as part of the digestion process; however, the stomach's digestive enzymes retain their activity under these conditions.

Protein folding is critical to its function. Scientists originally thought that the proteins themselves were responsible for the folding process. Only recently researchers discovered that often they receive assistance in the folding process from protein helpers, or **chaperones** (or chaperonins) that associate with the target protein during the folding process. They act by preventing polypeptide aggregation that comprise the complete protein structure, and they disassociate from the protein once the target protein is folded.

Link to Learning

For an additional perspective on proteins, view this animation called “Biomolecules: The Proteins.”

Section Summary

Proteins are a class of macromolecules that perform a diverse range of functions for the cell. They help in metabolism by acting as enzymes, carriers, or hormones, and provide structural support. The building blocks of proteins (monomers) are amino acids. Each amino acid has a central carbon that bonds to an amino group, a carboxyl group, a hydrogen atom, and an R group or side chain. There are 20 commonly occurring amino acids, each of which differs in the R group. A peptide bond links each amino acid to its neighbors. A long amino acid chain is a polypeptide.

Proteins are organized at four levels: primary, secondary, tertiary, and (optional) quaternary. The primary structure is the amino acids' unique sequence. The polypeptide's local folding to form structures such as the α-helix and β-pleated sheet constitutes the secondary structure. The overall three-dimensional structure is the tertiary structure. When two or more polypeptides combine to form

the complete protein structure, the configuration is the protein's quaternary structure. Protein shape and function are intricately linked. Any change in shape caused by changes in temperature or pH may lead to protein denaturation and a loss in function.

Visual Connection Questions

[link] Which categories of amino acid would you expect to find on the surface of a soluble protein, and which would you expect to find in the interior? What distribution of amino acids would you expect to find in a protein embedded in a lipid bilayer?

[link] Polar and charged amino acid residues (the remainder after peptide bond formation) are more likely to be found on the surface of soluble proteins where they can interact with water, and nonpolar (e.g., amino acid side chains) are more likely to be found in the interior where they are sequestered from water. In membrane proteins, nonpolar and hydrophobic amino acid side chains associate with the hydrophobic tails of phospholipids, while polar and charged amino acid side chains interact with the polar head groups or with the

aqueous solution. However, there are exceptions. Sometimes, positively and negatively charged amino acid side chains interact with one another in the interior of a protein, and polar or charged amino acid side chains that interact with a ligand can be found in the ligand binding pocket.

Review Questions

The monomers that make up proteins are called _____.

1. nucleotides
2. disaccharides
3. amino acids
4. chaperones

C

The α-helix and the β-pleated sheet are part of which protein structure?

1. primary
2. secondary
3. tertiary

4. quaternary

B

Mad cow disease is an infectious disease where one misfolded protein causes all other copies of the protein to begin misfolding. This is an example of a disease impacting ___ structure.

1. primary
2. secondary
3. tertiary
4. quaternary

C

Critical Thinking Questions

Explain what happens if even one amino acid is substituted for another in a polypeptide chain. Provide a specific example.

A change in gene sequence can lead to a different amino acid being added to a polypeptide chain instead of the normal one.

This causes a change in protein structure and function. For example, in sickle cell anemia, the hemoglobin β chain has a single amino acid substitution—the amino acid glutamic acid in position six is substituted by valine. Because of this change, hemoglobin molecules form aggregates, and the disc-shaped red blood cells assume a crescent shape, which results in serious health problems.

Describe the differences in the four protein structures.

The sequence and number of amino acids in a polypeptide chain is its primary structure. The local folding of the polypeptide in some regions is the secondary structure of the protein. The three-dimensional structure of a polypeptide is known as its tertiary structure, created in part by chemical interactions such as hydrogen bonds between polar side chains, van der Waals interactions, disulfide linkages, and hydrophobic interactions. Some proteins are formed from multiple polypeptides, also known as subunits, and the interaction of these subunits forms the quaternary structure.

Aquaporins are proteins embedded in the plasma membrane that allow water molecules

to move between the extracellular matrix and the intracellular space. Based on its function and location, describe the key features of the protein's shape and the chemical characteristics of its amino acids.

The protein must form a channel in the plasma membrane that allows water into the cell since water cannot cross the plasma membrane by itself. Since aquaporins are embedded in the plasma membrane and connect with both the intracellular and extracellular spaces, it must be amphipathic like the plasma membrane. The top and bottom of the protein must contain charged or polar amino acids (hydrophilic) to interact with the aqueous environments. The exterior transmembrane region must contain non-polar amino acids (hydrophobic) that can interact with the phospholipid tails. However, the inside of this channel must contain hydrophilic amino acids since they will interact with the traveling water molecules.

Glossary

alpha-helix structure (α-helix)
: type of secondary protein structure formed by folding the polypeptide into a helix shape with hydrogen bonds stabilizing the structure

amino acid

a protein's monomer; has a central carbon or alpha carbon to which an amino group, a carboxyl group, a hydrogen, and an R group or side chain is attached; the R group is different for all 20 common amino acids

beta-pleated sheet (β-pleated)

secondary structure in proteins in which hydrogen bonding forms “pleats” between atoms on the polypeptide chain's backbone

chaperone

(also, chaperonin) protein that helps nascent protein in the folding process

denaturation

loss of shape in a protein as a result of changes in temperature, pH, or chemical exposure

enzyme

catalyst in a biochemical reaction that is usually a complex or conjugated protein

hormone

chemical signaling molecule, usually protein or steroid, secreted by endocrine cells that act to control or regulate specific physiological processes

peptide bond

bond formed between two amino acids by a dehydration reaction

polypeptide
: long chain of amino acids that peptide bonds link

primary structure
: linear sequence of amino acids in a protein

protein
: biological macromolecule comprised of one or more amino acid chains

quaternary structure
: association of discrete polypeptide subunits in a protein

secondary structure
: regular structure that proteins form by intramolecular hydrogen bonding between the oxygen atom of one amino acid residue and the hydrogen attached to the nitrogen atom of another amino acid residue

tertiary structure
: a protein's three-dimensional conformation, including interactions between secondary structural elements; formed from interactions between amino acid side chains

Nucleic Acids
By the end of this section, you will be able to do the following:

- Describe nucleic acids' structure and define the two types of nucleic acids
- Explain DNA's structure and role
- Explain RNA's structure and roles

Nucleic acids are the most important macromolecules for the continuity of life. They carry the cell's genetic blueprint and carry instructions for its functioning.

Three components comprise a nucleotide: a nitrogenous base, a pentose sugar, and one or more phosphate groups. Carbon residues in the pentose are numbered 1′ through 5′ (the prime distinguishes these residues from those in the base, which are numbered without using a prime notation). The base is attached to the ribose's 1′ position, and the phosphate is attached to the 5′ position. When a polynucleotide forms, the incoming nucleotide's 5′ phosphate attaches to the 3′ hydroxyl group at the end of the growing chain. Two types of pentose are in nucleotides, deoxyribose (found in DNA) and ribose (found in RNA). Deoxyribose is similar in structure to ribose, but it has an H instead of an OH at the 2′ position. We can divide bases into two categories: purines and pyrimidines. Purines have a double ring structure, and pyrimidines have a single ring.

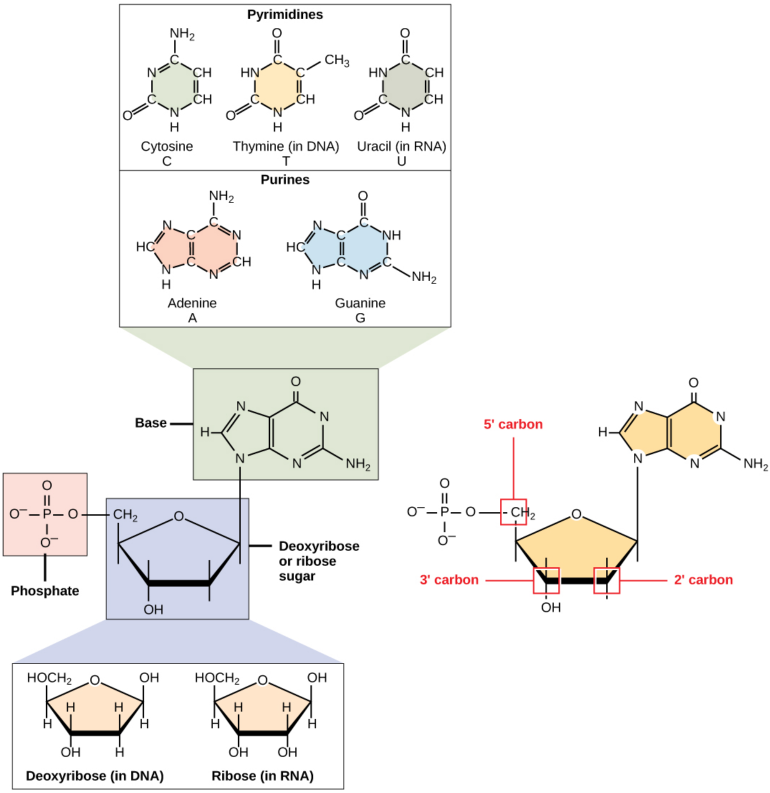
DNA and RNA

The two main types of nucleic acids are **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. DNA is the genetic material in all living organisms, ranging from single-celled bacteria to multicellular mammals. It is in the nucleus of eukaryotes and in the organelles, chloroplasts, and mitochondria. In prokaryotes, the DNA is not enclosed in a membranous envelope.

The cell's entire genetic content is its genome, and the study of genomes is genomics. In eukaryotic cells but not in prokaryotes, DNA forms a complex with histone proteins to form chromatin, the substance of eukaryotic chromosomes. A chromosome may contain tens of thousands of genes. Many genes contain the information to make protein products. Other genes code for RNA products. DNA controls all of the cellular activities by turning the genes "on" or "off."

The other type of nucleic acid, RNA, is mostly involved in protein synthesis. The DNA molecules never leave the nucleus but instead use an intermediary to communicate with the rest of the cell. This intermediary is the **messenger RNA (mRNA)**. Other types of RNA—like rRNA, tRNA, and microRNA—are involved in protein synthesis and its regulation.

DNA and RNA are comprised of monomers that scientists call **nucleotides**. The nucleotides combine with each other to form a **polynucleotide**, DNA or RNA. Three components comprise each nucleotide: a nitrogenous base, a pentose (five-carbon) sugar, and a phosphate group ([link]). Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to one or more phosphate groups.



The nitrogenous bases, important components of nucleotides, are organic molecules and are so named because they contain carbon and nitrogen. They are bases because they contain an amino group that has the potential of binding an extra hydrogen, and thus

decreasing the hydrogen ion concentration in its environment, making it more basic. Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G) cytosine (C), and thymine (T).

Scientists classify adenine and guanine as **purines**. The purine's primary structure is two carbon-nitrogen rings. Scientists classify cytosine, thymine, and uracil as **pyrimidines** which have a single carbon-nitrogen ring as their primary structure ([link]). Each of these basic carbon-nitrogen rings has different functional groups attached to it. In molecular biology shorthand, we know the nitrogenous bases by their symbols A, T, G, C, and U. DNA contains A, T, G, and C; whereas, RNA contains A, U, G, and C.

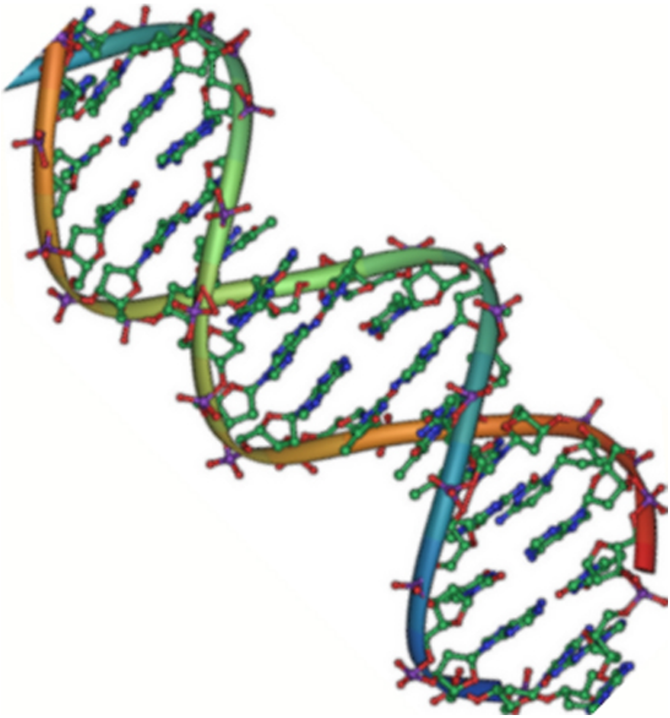
The pentose sugar in DNA is deoxyribose, and in RNA, the sugar is ribose ([link]). The difference between the sugars is the presence of the hydroxyl group on the ribose's second carbon and hydrogen on the deoxyribose's second carbon. The carbon atoms of the sugar molecule are numbered as 1′, 2′, 3′, 4′, and 5′ (1′ is read as “one prime”). The phosphate residue attaches to the hydroxyl group of the 5′ carbon of one sugar and the hydroxyl group of the 3′ carbon of the sugar of the next nucleotide, which forms a 5′–3′ **phosphodiester** linkage. A simple dehydration reaction like the other linkages connecting monomers in macromolecules does not

form the phosphodiester linkage. Its formation involves removing two phosphate groups. A polynucleotide may have thousands of such phosphodiester linkages.

Native DNA is an antiparallel double helix. The phosphate backbone (indicated by the curvy lines) is on the outside, and the bases are on the inside. Each base from one strand interacts via hydrogen bonding with a base from the opposing strand. (credit: Jerome Walker/Dennis Myts)

DNA Double-Helix Structure

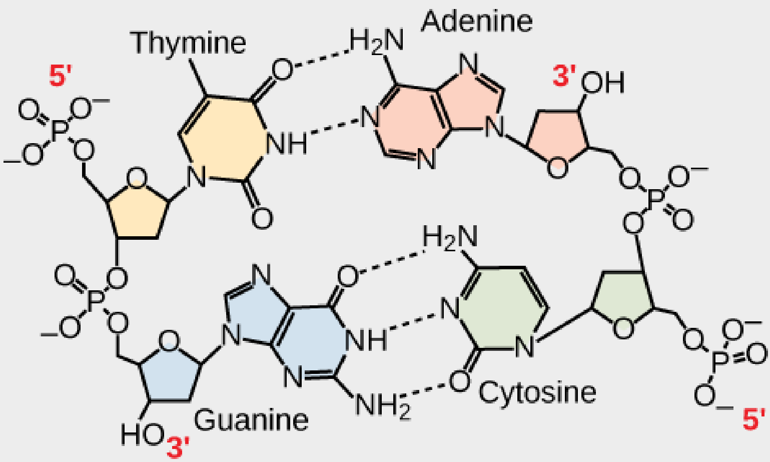
DNA has a double-helix structure ([link]). The sugar and phosphate lie on the outside of the helix, forming the DNA's backbone. The nitrogenous bases are stacked in the interior, like a pair of staircase steps. Hydrogen bonds bind the pairs to each other. Every base pair in the double helix is separated from the next base pair by 0.34 nm. The helix's two strands run in opposite directions, meaning that the 5′ carbon end of one strand will face the 3′ carbon end of its matching strand. (Scientists call this an antiparallel orientation and is important to DNA replication and in many nucleic acid interactions.)



Only certain types of base pairing are allowed. For example, a certain purine can only pair with a certain pyrimidine. This means A can pair with T, and G can pair with C, as [link] shows. This is the base complementary rule. In other words, the DNA strands are complementary to each other. If the sequence of one strand is AATTGGCC, the complementary strand would have the sequence TTAACCGG. During DNA replication, each strand copies itself, resulting in a daughter DNA double helix containing one parental DNA strand and a newly synthesized strand.

Visual Connection

In a double stranded DNA molecule, the two strands run antiparallel to one another so that one strand runs 5′ to 3′ and the other 3′ to 5′. The phosphate backbone is located on the outside, and the bases are in the middle. Adenine forms hydrogen bonds (or base pairs) with thymine, and guanine base pairs with cytosine.



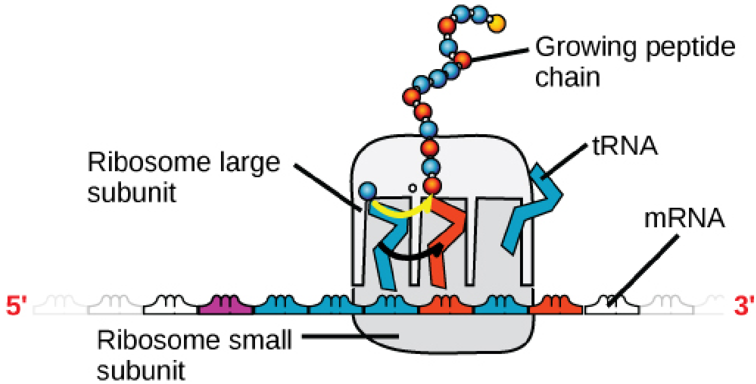
A mutation occurs, and adenine replaces cytosine. What impact do you think this will have on the DNA structure?

A ribosome has two parts: a large subunit and a small subunit. The mRNA sits in between the two subunits. A tRNA molecule recognizes a codon on the mRNA, binds to it by complementary base pairing, and adds the correct amino acid to the growing peptide chain.

RNA

Ribonucleic acid, or RNA, is mainly involved in the process of protein synthesis under the direction of DNA. RNA is usually single-stranded and is comprised of ribonucleotides that are linked by phosphodiester bonds. A ribonucleotide in the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and the phosphate group.

There are four major types of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and microRNA (miRNA). The first, mRNA, carries the message from DNA, which controls all of the cellular activities in a cell. If a cell requires synthesizing a certain protein, the gene for this product turns "on" and the messenger RNA synthesizes in the nucleus. The RNA base sequence is complementary to the DNA's coding sequence from which it has been copied. However, in RNA, the base T is absent and U is present instead. If the DNA strand has a sequence AATTGCGC, the sequence of the complementary RNA is UUAACGCG. In the cytoplasm, the mRNA interacts with ribosomes and other cellular machinery ([link]).



The mRNA is read in sets of three bases known as codons. Each codon codes for a single amino acid. In this way, the mRNA is read and the protein product is made. **Ribosomal RNA (rRNA)** is a major constituent of ribosomes on which the mRNA binds. The rRNA ensures the proper alignment of the mRNA and the Ribosomes. The ribosome's rRNA also has an enzymatic activity (peptidyl transferase) and catalyzes peptide bond formation between two aligned amino acids. **Transfer RNA (tRNA)** is one of the smallest of the four types of RNA, usually 70–90 nucleotides long. It carries the correct amino acid to the protein synthesis site. It is the base pairing between the tRNA and mRNA that allows for the correct amino acid to insert itself in the polypeptide chain. MicroRNAs are the smallest RNA molecules and their role involves regulating gene expression by interfering with the expression of certain mRNA messages. [link] summarizes DNA and RNA features.

DNA and RNA Features	DNA	RNA
Function	Carries genetic information	Involved in protein synthesis
Location	Remains in the nucleus	Leaves the nucleus
Structure	Double helix	Usually single-stranded
Sugar	Deoxyribose	Ribose
Pyrimidines	Cytosine, thymine	Cytosine, uracil
Purines	Adenine, guanine	Adenine, guanine

Even though the RNA is single stranded, most RNA types show extensive intramolecular base pairing between complementary sequences, creating a predictable three-dimensional structure essential for their function.

As you have learned, information flow in an organism takes place from DNA to RNA to protein. DNA dictates the structure of mRNA in a process scientists call **transcription**, and RNA dictates the protein's structure in a process scientists call **translation**. This is the Central Dogma of Life, which holds true for all organisms; however, exceptions to the rule occur in connection with viral infections.

Link to Learning
To learn more about DNA, explore the Howard Hughes Medical Institute BioInteractive animations on the topic of DNA.

Section Summary

Nucleic acids are molecules comprised of nucleotides that direct cellular activities such as cell division and protein synthesis. Pentose sugar, a nitrogenous base, and a phosphate group comprise each nucleotide. There are two types of nucleic acids: DNA and RNA. DNA carries the cell's genetic blueprint and passes it on from parents to offspring (in the form of chromosomes). It has a double-helical structure with the two strands running in opposite directions, connected by hydrogen bonds, and complementary to each other. RNA is a single-stranded polymer composed of linked nucleotides made up of a pentose sugar (ribose), a nitrogenous base, and a phosphate group. RNA is involved in protein synthesis and its regulation. Messenger RNA (mRNA) copies from the DNA, exports itself from the nucleus to the cytoplasm, and contains information for constructing proteins. Ribosomal RNA (rRNA) is a part of the ribosomes at the site of

protein synthesis; whereas, transfer RNA (tRNA) carries the amino acid to the site of protein synthesis. The microRNA regulates using mRNA for protein synthesis.

Visual Connection Questions

[link] A mutation occurs, and cytosine is replaced with adenine. What impact do you think this will have on the DNA structure?

[link] Adenine is larger than cytosine and will not be able to base pair properly with the guanine on the opposing strand. This will cause the DNA to bulge. DNA repair enzymes may recognize the bulge and replace the incorrect nucleotide.

Review Questions

A nucleotide of DNA may contain _____.

1. ribose, uracil, and a phosphate group
2. deoxyribose, uracil, and a phosphate group

3. deoxyribose, thymine, and a phosphate group
4. ribose, thymine, and a phosphate group

C

The building blocks of nucleic acids are ______.

1. sugars
2. nitrogenous bases
3. peptides
4. nucleotides

D

How does the double helix structure of DNA support its role in encoding the genome?

1. The sugar-phosphate backbone provides a template for DNA replication.
2. tRNA pairing with the template strand creates proteins encoded by the genome.
3. Complementary base pairing creates a very stable structure.
4. Complementary base pairing allows for easy editing of both strands of DNA.

C

Critical Thinking Questions

What are the structural differences between RNA and DNA?

DNA has a double-helix structure. The sugar and the phosphate are on the outside of the helix and the nitrogenous bases are in the interior. The monomers of DNA are nucleotides containing deoxyribose, one of the four nitrogenous bases (A, T, G and C), and a phosphate group. RNA is usually single-stranded and is made of ribonucleotides that are linked by phosphodiester linkages. A ribonucleotide contains ribose (the pentose sugar), one of the four nitrogenous bases (A,U, G, and C), and the phosphate group.

What are the four types of RNA and how do they function?

The four types of RNA are messenger RNA, ribosomal RNA, transfer RNA, and microRNA.

Messenger RNA carries the information from the DNA that controls all cellular activities. The mRNA binds to the ribosomes that are constructed of proteins and rRNA, and tRNA transfers the correct amino acid to the site of protein synthesis. microRNA regulates the availability of mRNA for translation.

Glossary

deoxyribonucleic acid (DNA)
: double-helical molecule that carries the cell's hereditary information

messenger RNA (mRNA)
: RNA that carries information from DNA to ribosomes during protein synthesis

nucleic acid
: biological macromolecule that carries the cell's genetic blueprint and carries instructions for the cell's functioning

nucleotide
: monomer of nucleic acids; contains a pentose sugar, one or more phosphate groups, and a nitrogenous base

phosphodiester
: **linkage** covalent chemical bond that holds together the polynucleotide chains with a

phosphate group linking neighboring nucleotides' two pentose sugars

polynucleotide
: long chain of nucleotides

purine
: type of nitrogenous base in DNA and RNA; adenine and guanine are purines

pyrimidine
: type of nitrogenous base in DNA and RNA; cytosine, thymine, and uracil are pyrimidines

ribonucleic acid (RNA)
: single-stranded, often internally base paired, molecule that is involved in protein synthesis

ribosomal RNA (rRNA)
: RNA that ensures the proper alignment of the mRNA and the ribosomes during protein synthesis and catalyzes forming the peptide linkage

transcription
: process through which messenger RNA forms on a template of DNA

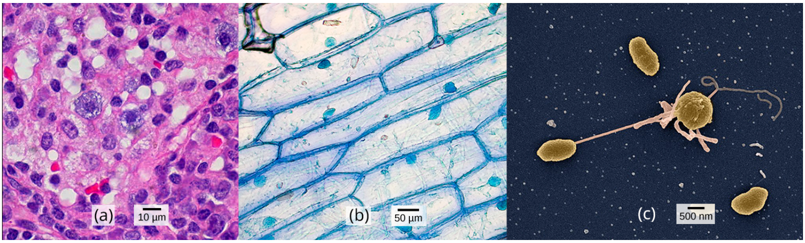
transfer RNA (tRNA)
: RNA that carries activated amino acids to the site of protein synthesis on the ribosome

translation

process through which RNA directs the protein's formation

Introduction

class="introduction" (a) Nasal sinus cells (viewed with a light microscope), (b) onion cells (viewed with a light microscope), and (c) *Vibrio tasmaniensis* bacterial cells (seen through a scanning electron microscope) are from very different organisms, yet all share certain basic cell structure characteristics. (credit a: modification of work by Ed Uthman, MD; credit b: modification of work by Umberto Salvagnin; credit c: modification of work by Anthony D'Onofrio, William H. Fowle, Eric J. Stewart, and Kim Lewis of the Lewis Lab at Northeastern University; scale-bar data from Matt Russell)



Close your eyes and picture a brick wall. What is the wall's basic building block? It is a single brick. Like a brick wall, cells are the building blocks that make up your body.

Your body has many kinds of cells, each specialized for a specific purpose. Just as we use a variety of materials to build a home, the human body is constructed from many cell types. For example, epithelial cells protect the body's surface and cover

the organs and body cavities within. Bone cells help to support and protect the body. Immune system cells fight invading bacteria. Additionally, blood and blood cells carry nutrients and oxygen throughout the body while removing carbon dioxide. Each of these cell types plays a vital role during the body's growth, development, and day-to-day maintenance. In spite of their enormous variety, however, cells from all organisms—even ones as diverse as bacteria, onion, and human—share certain fundamental characteristics.

Studying Cells

By the end of this section, you will be able to do the following:

- Describe the role of cells in organisms
- Compare and contrast light microscopy and electron microscopy
- Summarize cell theory

A cell is the smallest unit of a living thing. Whether comprised of one cell (like bacteria) or many cells (like a human), we call it an organism. Thus, cells are the basic building blocks of all organisms.

Several cells of one kind that interconnect with each other and perform a shared function form tissues. These tissues combine to form an organ (your stomach, heart, or brain), and several organs comprise an organ system (such as the digestive system, circulatory system, or nervous system). Several systems that function together form an organism (like a human being). Here, we will examine the structure and function of cells.

There are many types of cells, which scientists group into one of two broad categories: prokaryotic and eukaryotic. For example, we classify both animal and plant cells as eukaryotic cells; whereas, we classify bacterial cells as prokaryotic. Before discussing the criteria for determining whether a cell is prokaryotic or eukaryotic, we will first

examine how biologists study cells.
(a) Most light microscopes in a college biology lab can magnify cells up to approximately 400 times and have a resolution of about 200 nanometers. (b) Electron microscopes provide a much higher magnification, 100,000x, and a have a resolution of 50 picometers. (credit a: modification of work by "GcG"/Wikimedia Commons; credit b: modification of work by Evan Bench) (a) These *Salmonella* bacteria appear as tiny purple dots when viewed with a light microscope. (b) This scanning electron microscope micrograph shows *Salmonella* bacteria (in red) invading human cells (yellow). Even though subfigure (b) shows a different *Salmonella* specimen than subfigure (a), you can still observe the comparative increase in magnification and detail. (credit a: modification of work by CDC/Armed Forces Institute of Pathology, Charles N. Farmer, Rocky Mountain Laboratories; credit b: modification of work by NIAID, NIH; scale-bar data from Matt Russell)

Microscopy

Cells vary in size. With few exceptions, we cannot see individual cells with the naked eye, so scientists use microscopes (micro- = “small”; -scope = “to look at”) to study them. A **microscope** is an instrument that magnifies an object. We photograph most cells with a microscope, so we can call these images micrographs.

The optics of a microscope's lenses change the image orientation that the user sees. A specimen that is right-side up and facing right on the microscope slide will appear upside-down and facing left when one views through a microscope, and vice versa. Similarly, if one moves the slide left while looking through the microscope, it will appear to move right, and if one moves it down, it will seem to move up. This occurs because microscopes use two sets of lenses to magnify the image. Because of the manner by which light travels through the lenses, this two lens system produces an inverted image (binocular, or dissecting microscopes, work in a similar manner, but include an additional magnification system that makes the final image appear to be upright).

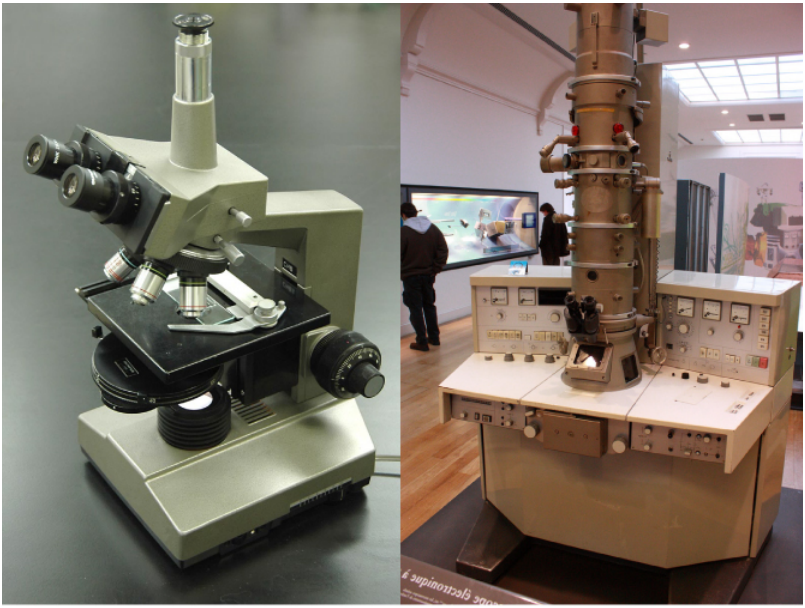
Light Microscopes

To give you a sense of cell size, a typical human red blood cell is about eight millionths of a meter or eight micrometers (abbreviated as eight μm) in diameter. A pin head is about two thousandths of a meter (two mm) in diameter. That means about 250 red blood cells could fit on a pinhead.

Most student microscopes are **light microscopes** ([link]**a**). Visible light passes and bends through the lens system to enable the user to see the specimen. Light microscopes are advantageous for viewing living organisms, but since individual cells are

generally transparent, their components are not distinguishable unless they are colored with special stains. Staining, however, usually kills the cells.

Light microscopes that undergraduates commonly use in the laboratory magnify up to approximately 400 times. Two parameters that are important in microscopy are magnification and resolving power. Magnification is the process of enlarging an object in appearance. Resolving power is the microscope's ability to distinguish two adjacent structures as separate: the higher the resolution, the better the image's clarity and detail. When one uses oil immersion lenses to study small objects, magnification usually increases to 1,000 times. In order to gain a better understanding of cellular structure and function, scientists typically use electron microscopes.



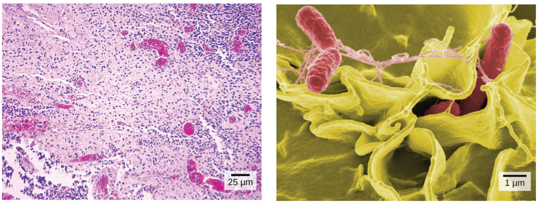
(a) (b)

Electron Microscopes

In contrast to light microscopes, **electron microscopes** ([link]**b**) use a beam of electrons instead of a beam of light. Not only does this allow for higher magnification and, thus, more detail ([link]), it also provides higher resolving power. The method to prepare the specimen for viewing with an electron microscope kills the specimen. Electrons have short wavelengths (shorter than photons) that move best in a vacuum, so we cannot view living cells with an electron microscope.

In a scanning electron microscope, a beam of electrons moves back and forth across a cell's

surface, creating details of cell surface characteristics. In a transmission electron microscope, the electron beam penetrates the cell and provides details of a cell’s internal structures. As you might imagine, electron microscopes are significantly more bulky and expensive than light microscopes.



(a) (b)

> **Link to Learning**
> For another perspective on cell size, try the HowBig interactive at this site.

Cell Theory

The microscopes we use today are far more complex than those that Dutch shopkeeper Antony van Leeuwenhoek, used in the 1600s. Skilled in crafting lenses, van Leeuwenhoek observed the movements of single-celled organisms, which he collectively termed “animalcules.”

In the 1665 publication *Micrographia*, experimental scientist Robert Hooke coined the term “cell” for the box-like structures he observed when viewing cork tissue through a lens. In the 1670s, van Leeuwenhoek discovered bacteria and protozoa. Later advances in lenses, microscope construction, and staining techniques enabled other scientists to see some components inside cells.

By the late 1830s, botanist Matthias Schleiden and zoologist Theodor Schwann were studying tissues and proposed the **unified cell theory**, which states that one or more cells comprise all living things, the cell is the basic unit of life, and new cells arise from existing cells. Rudolf Virchow later made important contributions to this theory.

Career Connection

Cytotechnologist

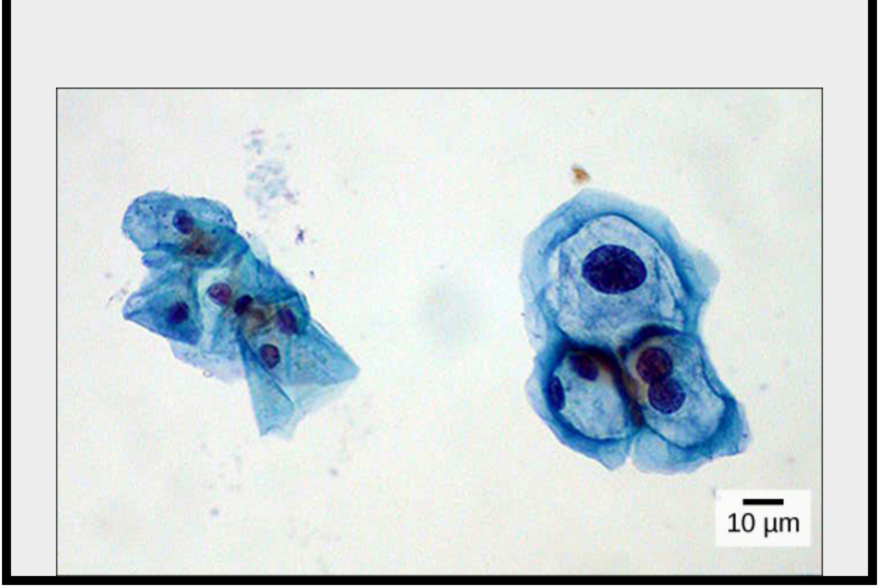
Have you ever heard of a medical test called a Pap smear ([link])? In this test, a doctor takes a small sample of cells from the patient's uterine cervix and sends it to a medical lab where a cytotechnologist stains the cells and examines them for any changes that could indicate cervical cancer or a microbial infection.

Cytotechnologists (cyto- = “cell”) are professionals who study cells via microscopic examinations and other laboratory tests. They are trained to

determine which cellular changes are within normal limits and which are abnormal. Their focus is not limited to cervical cells. They study cellular specimens that come from all organs. When they notice abnormalities, they consult a pathologist, a medical doctor who interprets and diagnoses changes that disease in body tissue and fluids cause.

Cytotechnologists play a vital role in saving people's lives. When doctors discover abnormalities early, a patient's treatment can begin sooner, which usually increases the chances of a successful outcome.

These uterine cervix cells, viewed through a light microscope, are from a Pap smear. Normal cells are on the left. The cells on the right are infected with human papillomavirus (HPV). Notice that the infected cells are larger. Also, two of these cells each have two nuclei instead of one, the normal number. (credit: modification of work by Ed Uthman, MD; scale-bar data from Matt Russell)



Section Summary

A cell is the smallest unit of life. Most cells are so tiny that we cannot see them with the naked eye. Therefore, scientists use microscopes to study cells. Electron microscopes provide higher magnification, higher resolution, and more detail than light microscopes. The unified cell theory states that one or more cells comprise all organisms, the cell is the basic unit of life, and new cells arise from existing cells.

Review Questions

When viewing a specimen through a light microscope, scientists use ______ to distinguish the individual components of cells.

1. a beam of electrons
2. radioactive isotopes
3. special stains
4. high temperatures

C

The ______ is the basic unit of life.

1. organism
2. cell
3. tissue
4. organ

B

Critical Thinking Questions

In your everyday life, you have probably noticed that certain instruments are ideal for

certain situations. For example, you would use a spoon rather than a fork to eat soup because a spoon is shaped for scooping, while soup would slip between the tines of a fork. The use of ideal instruments also applies in science. In what situation(s) would the use of a light microscope be ideal, and why?

A light microscope would be ideal when viewing a small living organism, especially when the cell has been stained to reveal details.

In what situation(s) would the use of a scanning electron microscope be ideal, and why?

A scanning electron microscope would be ideal when you want to view the minute details of a cell's surface, because its beam of electrons moves back and forth over the surface to convey the image.

In what situation(s) would a transmission electron microscope be ideal, and why?

A transmission electron microscope would be ideal for viewing the cell's internal structures, because many of the internal structures have

membranes that are not visible by the light microscope.

What are the advantages and disadvantages of each of these types of microscopes?

The advantages of light microscopes are that they are easily obtained, and the light beam does not kill the cells. However, typical light microscopes are somewhat limited in the amount of detail they can reveal. Electron microscopes are ideal because you can view intricate details, but they are bulky and costly, and preparation for the microscopic examination kills the specimen.

Explain how the formation of an adult human follows the cell theory.

The cell theory states:

1. All living things are made of cells.;
2. Cells are the most basic unit of life.;
3. New cells arise from existing cells.

All humans are multicellular organisms whose smallest building blocks are cells. Adult humans begin with the fusion of a male gamete cell

with a female gamete cell to form a fertilized egg (single cell). That cell then divides into two cells, which each divides into two more cells, and so forth until all the cells of a human embryo are made. As the embryo passes through all the developmental stages to make an adult human, the cells that are added arise from division of existing cells.

Glossary

cell theory
: see unified cell theory

electron microscope
: an instrument that magnifies an object using an electron beam that passes and bends through a lens system to visualize a specimen

light microscope
: an instrument that magnifies an object using a beam of visible light that passes and bends through a lens system to visualize a specimen

microscope
: an instrument that magnifies an object

unified cell theory
: a biological concept that states that one or more cells comprise all organisms; the cell is the basic unit of life; and new cells arise from

existing cells

Prokaryotic Cells
By the end of this section, you will be able to do the following:

- Name examples of prokaryotic and eukaryotic organisms
- Compare and contrast prokaryotic and eukaryotic cells
- Describe the relative sizes of different cells
- Explain why cells must be small

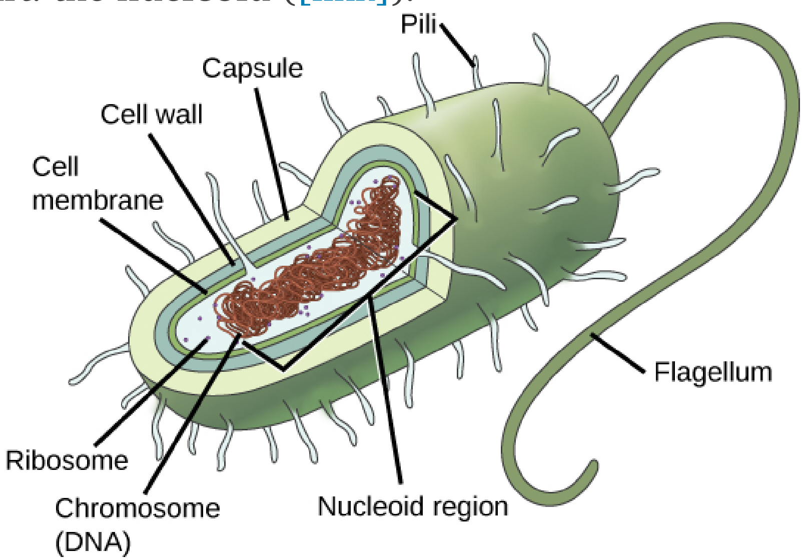
Cells fall into one of two broad categories: prokaryotic and eukaryotic. We classify only the predominantly single-celled organisms Bacteria and Archaea as prokaryotes (pro- = "before"; -kary- = "nucleus"). Animal cells, plants, fungi, and protists are all eukaryotes (eu- = "true").
This figure shows the generalized structure of a prokaryotic cell. All prokaryotes have chromosomal DNA localized in a nucleoid, ribosomes, a cell membrane, and a cell wall. The other structures shown are present in some, but not all, bacteria.
This figure shows relative sizes of microbes on a logarithmic scale (recall that each unit of increase in a logarithmic scale represents a 10-fold increase in the quantity measured).

Components of Prokaryotic Cells

All cells share four common components: 1) a

plasma membrane, an outer covering that separates the cell's interior from its surrounding environment; 2) cytoplasm, consisting of a jelly-like cytosol within the cell in which there are other cellular components; 3) DNA, the cell's genetic material; and 4) ribosomes, which synthesize proteins. However, prokaryotes differ from eukaryotic cells in several ways.

A **prokaryote** is a simple, mostly single-celled (unicellular) organism that lacks a nucleus, or any other membrane-bound organelle. We will shortly come to see that this is significantly different in eukaryotes. Prokaryotic DNA is in the cell's central part: the **nucleoid** ([link]).



Most prokaryotes have a peptidoglycan cell wall and many have a polysaccharide capsule ([link]). The cell wall acts as an extra layer of protection, helps

the cell maintain its shape, and prevents dehydration. The capsule enables the cell to attach to surfaces in its environment. Some prokaryotes have flagella, pili, or fimbriae. Flagella are used for locomotion. Pili exchange genetic material during conjugation, the process by which one bacterium transfers genetic material to another through direct contact. Bacteria use fimbriae to attach to a host cell.

Career Connection

Microbiologist

The most effective action anyone can take to prevent the spread of contagious illnesses is to wash his or her hands. Why? Because microbes (organisms so tiny that they can only be seen with microscopes) are ubiquitous. They live on doorknobs, money, your hands, and many other surfaces. If someone sneezes into his hand and touches a doorknob, and afterwards you touch that same doorknob, the microbes from the sneezer's mucus are now on your hands. If you touch your hands to your mouth, nose, or eyes, those microbes can enter your body and could make you sick. However, not all microbes (also called microorganisms) cause disease; most are actually beneficial. You have microbes in your gut that make vitamin K. Other microorganisms are used to ferment beer and wine.

Microbiologists are scientists who study microbes. Microbiologists can pursue a number of careers. Not only do they work in the food industry, they are also employed in the veterinary and medical fields. They can work in the pharmaceutical sector, serving key roles in research and development by identifying new antibiotic sources that can treat bacterial infections.

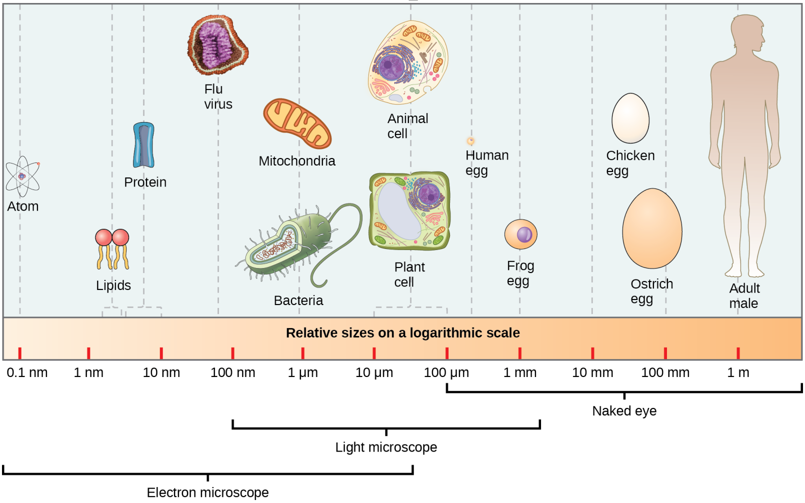
Environmental microbiologists may look for new ways to use specially selected or genetically engineered microbes to remove pollutants from soil or groundwater, as well as hazardous elements from contaminated sites. We call using these microbes bioremediation technologies.

Microbiologists can also work in the bioinformatics field, providing specialized knowledge and insight for designing, developing, and specificity of computer models of, for example, bacterial epidemics.

Cell Size

At 0.1 to 5.0 μm in diameter, prokaryotic cells are significantly smaller than eukaryotic cells, which have diameters ranging from 10 to 100 μm ([link]). The prokaryotes' small size allows ions and organic molecules that enter them to quickly diffuse to other parts of the cell. Similarly, any wastes produced within a prokaryotic cell can quickly diffuse. This is

not the case in eukaryotic cells, which have developed different structural adaptations to enhance intracellular transport.

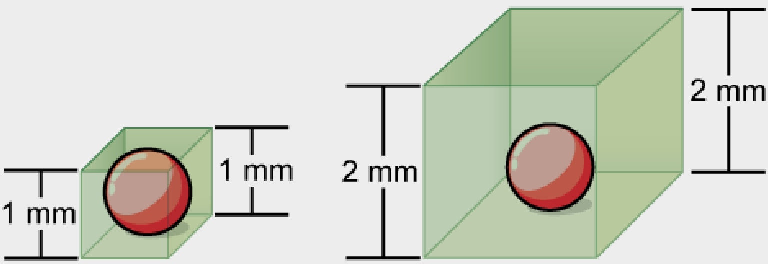


Small size, in general, is necessary for all cells, whether prokaryotic or eukaryotic. Let's examine why that is so. First, we'll consider the area and volume of a typical cell. Not all cells are spherical in shape, but most tend to approximate a sphere. You may remember from your high school geometry course that the formula for the surface area of a sphere is $4\pi r^2$, while the formula for its volume is $4\pi r^3/3$. Thus, as the radius of a cell increases, its surface area increases as the square of its radius, but its volume increases as the cube of its radius (much more rapidly). Therefore, as a cell increases in size, its surface area-to-volume ratio decreases. This same principle would apply if the cell had a cube shape ([link]). If the cell grows too large, the plasma membrane will not have sufficient surface area to

support the rate of diffusion required for the increased volume. In other words, as a cell grows, it becomes less efficient. One way to become more efficient is to divide. Another way is to develop organelles that perform specific tasks. These adaptations lead to developing more sophisticated cells, which we call eukaryotic cells.

Visual Connection

Notice that as a cell increases in size, its surface area-to-volume ratio decreases. When there is insufficient surface area to support a cell's increasing volume, a cell will either divide or die. The cell on the left has a volume of 1 mm3 and a surface area of 6 mm2, with a surface area-to-volume ratio of 6 to 1; whereas, the cell on the right has a volume of 8 mm3 and a surface area of 24 mm2, with a surface area-to-volume ratio of 3 to 1.



Prokaryotic cells are much smaller than eukaryotic cells. What advantages might small cell size confer on a cell? What advantages might large cell size have?

Section Summary

Prokaryotes are single-celled organisms of the domains Bacteria and Archaea. All prokaryotes have plasma membranes, cytoplasm, ribosomes, and DNA that is not membrane-bound. Most have peptidoglycan cell walls and many have polysaccharide capsules. Prokaryotic cells range in diameter from 0.1 to 5.0 μm.

As a cell increases in size, its surface area-to-volume ratio decreases. If the cell grows too large, the plasma membrane will not have sufficient surface area to support the rate of diffusion required for the increased volume.

Visual Connection Questions

[link] Prokaryotic cells are much smaller than eukaryotic cells. What advantages might small cell size confer on a cell? What advantages might large cell size have?

[link] Substances can diffuse more quickly through small cells. Small cells have no need

for organelles and therefore do not need to expend energy getting substances across organelle membranes. Large cells have organelles that can separate cellular processes, enabling them to build molecules that are more complex.

Review Questions

Prokaryotes depend on ______ to obtain some materials and to get rid of wastes.

1. ribosomes
2. flagella
3. cell division
4. diffusion

D

Bacteria that lack fimbriae are less likely to ______.

1. adhere to cell surfaces
2. swim through bodily fluids
3. synthesize proteins
4. retain the ability to divide

A

Which of the following organisms is a prokaryote?

1. amoeba
2. influenza A virus
3. charophyte algae
4. *E. coli*

D

Critical Thinking Questions

Antibiotics are medicines that are used to fight bacterial infections. These medicines kill prokaryotic cells without harming human cells. What part or parts of the bacterial cell do you think antibiotics target? Why?

The cell wall would be targeted by antibiotics as well as the bacteria’s ability to replicate. This would inhibit the bacteria’s ability to reproduce, and it would compromise its defense mechanisms.

Explain why not all microbes are harmful.

Some microbes are beneficial. For instance, *E. coli* bacteria populate the human gut and help break down fiber in the diet. Some foods such as yogurt are formed by bacteria.

Glossary

nucleoid
: central part of a prokaryotic cell's central part where the chromosome is located

prokaryote
: unicellular organism that lacks a nucleus or any other membrane-bound organelle

Eukaryotic Cells

By the end of this section, you will be able to do the following:

- Describe the structure of eukaryotic cells
- Compare animal cells with plant cells
- State the role of the plasma membrane
- Summarize the functions of the major cell organelles

Have you ever heard the phrase "form follows function?" It's a philosophy that many industries follow. In architecture, this means that buildings should be constructed to support the activities that will be carried out inside them. For example, a skyscraper should include several elevator banks. A hospital should have its emergency room easily accessible.

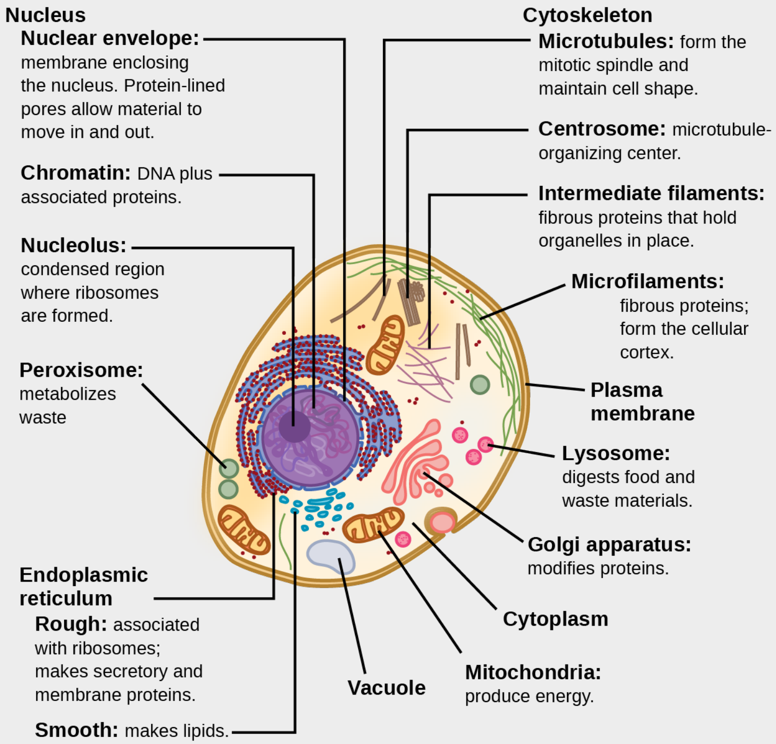
Our natural world also utilizes the principle of form following function, especially in cell biology, and this will become clear as we explore eukaryotic cells ([link]). Unlike prokaryotic cells, **eukaryotic cells** have: 1) a membrane-bound nucleus; 2) numerous membrane-bound **organelles** such as the endoplasmic reticulum, Golgi apparatus, chloroplasts, mitochondria, and others; and 3) several, rod-shaped chromosomes. Because a membrane surrounds eukaryotic cell's nucleus, it has a "true nucleus." The word "organelle" means "little organ," and, as we already mentioned,

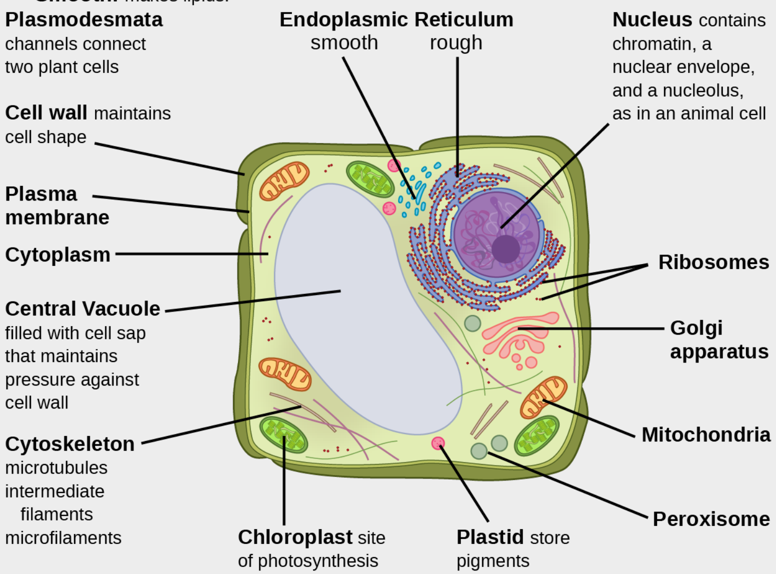
organelles have specialized cellular functions, just as your body's organs have specialized functions.

At this point, it should be clear to you that eukaryotic cells have a more complex structure than prokaryotic cells. Organelles allow different functions to be compartmentalized in different areas of the cell. Before turning to organelles, let’s first examine two important components of the cell: the plasma membrane and the cytoplasm.

Visual Connection

These figures show the major organelles and other cell components of (a) a typical animal cell and (b) a typical eukaryotic plant cell. The plant cell has a cell wall, chloroplasts, plastids, and a central vacuole—structures not in animal cells. Most cells do not have lysosomes or centrosomes.





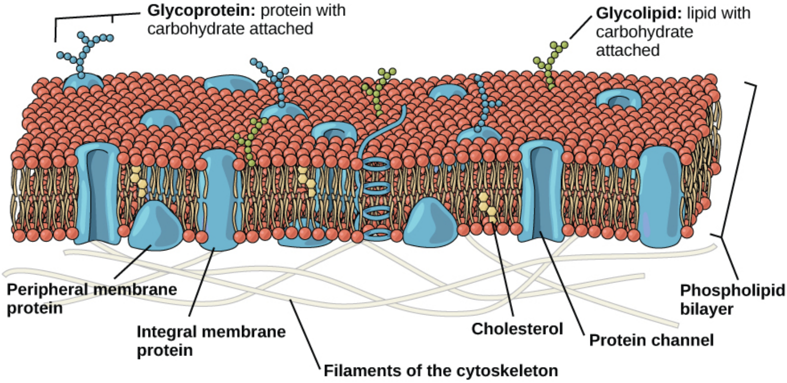
If the nucleolus were not able to carry out its function, what other cellular organelles would be

affected?

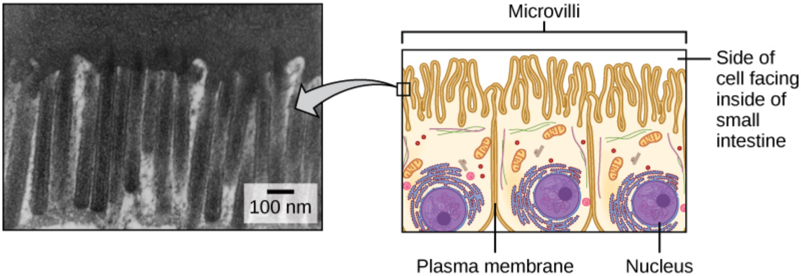
The eukaryotic plasma membrane is a phospholipid bilayer with proteins and cholesterol embedded in it. Microvilli, as they appear on cells lining the small intestine, increase the surface area available for absorption. These microvilli are only on the area of the plasma membrane that faces the cavity from which substances will be absorbed. (credit "micrograph": modification of work by Louisa Howard)

The Plasma Membrane

Like prokaryotes, eukaryotic cells have a **plasma membrane** ([link]), a phospholipid bilayer with embedded proteins that separates the internal contents of the cell from its surrounding environment. A phospholipid is a lipid molecule with two fatty acid chains and a phosphate-containing group. The plasma membrane controls the passage of organic molecules, ions, water, and oxygen into and out of the cell. Wastes (such as carbon dioxide and ammonia) also leave the cell by passing through the plasma membrane.



The plasma membranes of cells that specialize in absorption fold into fingerlike projections that we call microvilli (singular = microvillus); ([link]). Such cells typically line the small intestine, the organ that absorbs nutrients from digested food. This is an excellent example of form following function. People with celiac disease have an immune response to gluten, which is a protein in wheat, barley, and rye. The immune response damages microvilli, and thus, afflicted individuals cannot absorb nutrients. This leads to malnutrition, cramping, and diarrhea. Patients suffering from celiac disease must follow a gluten-free diet.



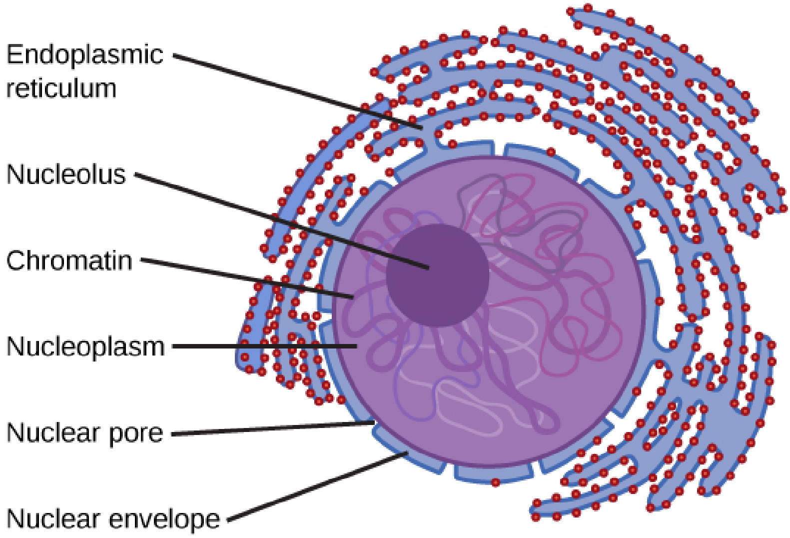
The Cytoplasm

The **cytoplasm** is the cell's entire region between the plasma membrane and the nuclear envelope (a structure we will discuss shortly). It is comprised of organelles suspended in the gel-like **cytosol**, the cytoskeleton, and various chemicals ([link]). Even though the cytoplasm consists of 70 to 80 percent water, it has a semi-solid consistency, which comes from the proteins within it. However, proteins are not the only organic molecules in the cytoplasm. Glucose and other simple sugars, polysaccharides, amino acids, nucleic acids, fatty acids, and derivatives of glycerol are also there. Ions of sodium, potassium, calcium, and many other elements also dissolve in the cytoplasm. Many metabolic reactions, including protein synthesis, take place in the cytoplasm.

The nucleus stores chromatin (DNA plus proteins) in a gel-like substance called the nucleoplasm. The nucleolus is a condensed chromatin region where ribosome synthesis occurs. We call the nucleus' boundary the nuclear envelope. It consists of two phospholipid bilayers: an outer and an inner membrane. The nuclear membrane is continuous with the endoplasmic reticulum. Nuclear pores allow substances to enter and exit the nucleus. (a) This image shows various levels of chromatin's organization (DNA and protein). (b) This image shows paired chromosomes. (credit b: modification of work by NIH; scale-bar data from Matt Russell)

The Nucleus

Typically, the nucleus is the most prominent organelle in a cell ([link]). The **nucleus** (plural = nuclei) houses the cell’s DNA and directs the synthesis of ribosomes and proteins. Let’s look at it in more detail ([link]).



The Nuclear Envelope

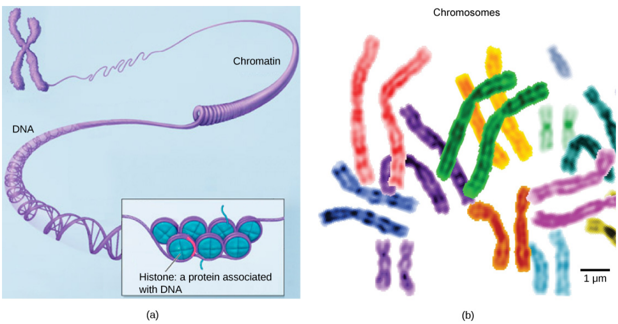
The **nuclear envelope** is a double-membrane structure that constitutes the nucleus' outermost portion ([link]). Both the nuclear envelope's inner and outer membranes are phospholipid bilayers.

The nuclear envelope is punctuated with pores that control the passage of ions, molecules, and RNA between the nucleoplasm and cytoplasm. The **nucleoplasm** is the semi-solid fluid inside the

nucleus, where we find the chromatin and the nucleolus.

Chromatin and Chromosomes

To understand chromatin, it is helpful to first explore **chromosomes**, structures within the nucleus that are made up of DNA, the hereditary material. You may remember that in prokaryotes, DNA is organized into a single circular chromosome. In eukaryotes, chromosomes are linear structures. Every eukaryotic species has a specific number of chromosomes in the nucleus of each cell. For example, in humans, the chromosome number is 46, while in fruit flies, it is eight. Chromosomes are only visible and distinguishable from one another when the cell is getting ready to divide. When the cell is in the growth and maintenance phases of its life cycle, proteins attach to chromosomes, and they resemble an unwound, jumbled bunch of threads. We call these unwound protein-chromosome complexes **chromatin** ([link]). Chromatin describes the material that makes up the chromosomes both when condensed and decondensed.



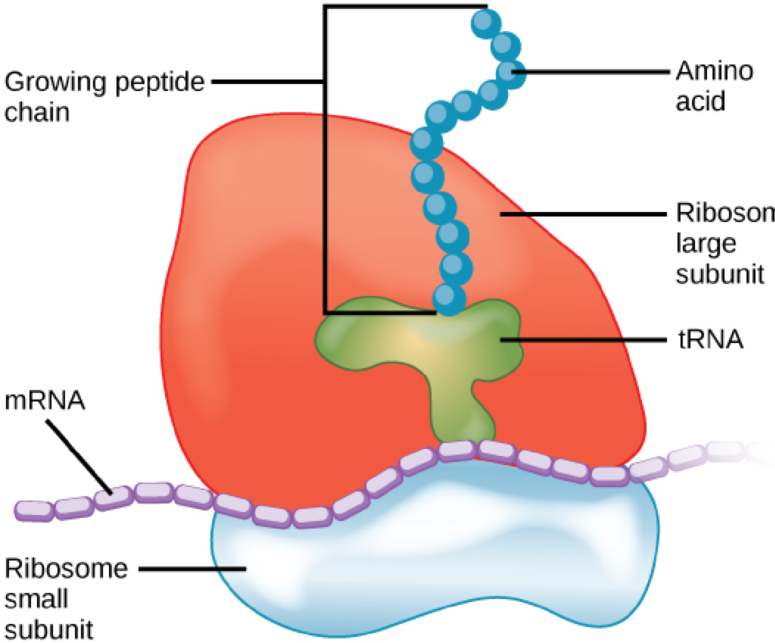
The Nucleolus

We already know that the nucleus directs the synthesis of ribosomes, but how does it do this? Some chromosomes have sections of DNA that encode ribosomal RNA. A darkly staining area within the nucleus called the **nucleolus** (plural = nucleoli) aggregates the ribosomal RNA with associated proteins to assemble the ribosomal subunits that are then transported out through the pores in the nuclear envelope to the cytoplasm. A large subunit (top) and a small subunit (bottom) comprise ribosomes. During protein synthesis, ribosomes assemble amino acids into proteins.

Ribosomes

Ribosomes are the cellular structures responsible for protein synthesis. When we view them through an electron microscope, ribosomes appear either as clusters (polyribosomes) or single, tiny dots that float freely in the cytoplasm. They may be attached to the plasma membrane's cytoplasmic side or the endoplasmic reticulum's cytoplasmic side and the nuclear envelope's outer membrane ([link]). Electron microscopy shows us that ribosomes, which are large protein and RNA complexes, consist of two subunits, large and small ([link]). Ribosomes receive their “orders” for protein synthesis from the nucleus where the DNA transcribes into messenger RNA (mRNA). The mRNA travels to the ribosomes, which

translate the code provided by the sequence of the nitrogenous bases in the mRNA into a specific order of amino acids in a protein. Amino acids are the building blocks of proteins.



Because protein synthesis is an essential function of all cells (including enzymes, hormones, antibodies, pigments, structural components, and surface receptors), there are ribosomes in practically every cell. Ribosomes are particularly abundant in cells that synthesize large amounts of protein. For example, the pancreas is responsible for creating several digestive enzymes and the cells that produce these enzymes contain many ribosomes. Thus, we see another example of form following function.

This electron micrograph shows a mitochondrion through an electron microscope. This organelle has

an outer membrane and an inner membrane. The inner membrane contains folds, called cristae, which increase its surface area. We call the space between the two membranes the intermembrane space, and the space inside the inner membrane the mitochondrial matrix. ATP synthesis takes place on the inner membrane. (credit: modification of work by Matthew Britton; scale-bar data from Matt Russell)

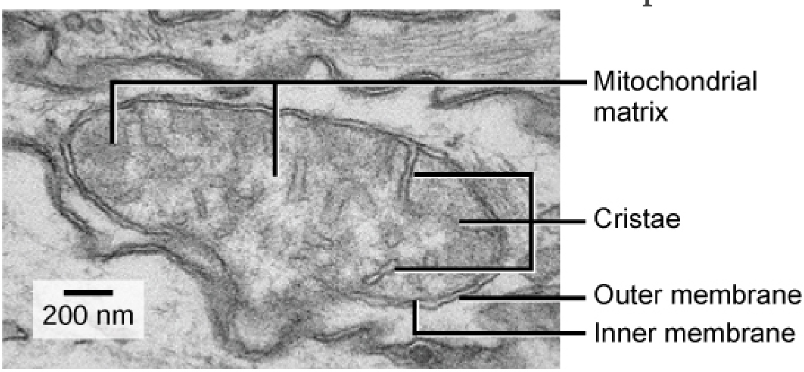
Mitochondria

Scientists often call **mitochondria** (singular = mitochondrion) the cell's “powerhouses” or “energy factories” because they are responsible for making adenosine triphosphate (ATP), the cell’s main energy-carrying molecule. ATP represents the cell's short-term stored energy. Cellular respiration is the process of making ATP using the chemical energy in glucose and other nutrients. In mitochondria, this process uses oxygen and produces carbon dioxide as a waste product. In fact, the carbon dioxide that you exhale with every breath comes from the cellular reactions that produce carbon dioxide as a byproduct.

In keeping with our theme of form following function, it is important to point out that muscle cells have a very high concentration of mitochondria that produce ATP. Your muscle cells need considerable energy to keep your body

moving. When your cells don't get enough oxygen, they do not make much ATP. Instead, producing lactic acid accompanies the small amount of ATP they make in the absence of oxygen.

Mitochondria are oval-shaped, double membrane organelles ([link]) that have their own ribosomes and DNA. Each membrane is a phospholipid bilayer embedded with proteins. The inner layer has folds called cristae. We call the area surrounded by the folds the mitochondrial matrix. The cristae and the matrix have different roles in cellular respiration.



Peroxisomes

Peroxisomes are small, round organelles enclosed by single membranes. They carry out oxidation reactions that break down fatty acids and amino acids. They also detoxify many poisons that may enter the body. (Many of these oxidation reactions release hydrogen peroxide, H_2O_2, which would be damaging to cells; however, when these reactions

are confined to peroxisomes, enzymes safely break down the H_2O_2 into oxygen and water.) For example, peroxisomes in liver cells detoxify alcohol. Glyoxysomes, which are specialized peroxisomes in plants, are responsible for converting stored fats into sugars. Plant cells contain many different types of peroxisomes that play a role in metabolism, pathogene defense, and stress response, to mention a few.

Vesicles and Vacuoles

Vesicles and **vacuoles** are membrane-bound sacs that function in storage and transport. Other than the fact that vacuoles are somewhat larger than vesicles, there is a very subtle distinction between them. Vesicle membranes can fuse with either the plasma membrane or other membrane systems within the cell. Additionally, some agents such as enzymes within plant vacuoles break down macromolecules. The vacuole's membrane does not fuse with the membranes of other cellular components.

The centrosome consists of two centrioles that lie at right angles to each other. Each centriole is a cylinder comprised of nine triplets of microtubules. Nontubulin proteins (indicated by the green lines) hold the microtubule triplets together. Cellulose is a long chain of β-glucose molecules connected by a 1-4 linkage. The dashed lines at each end of the

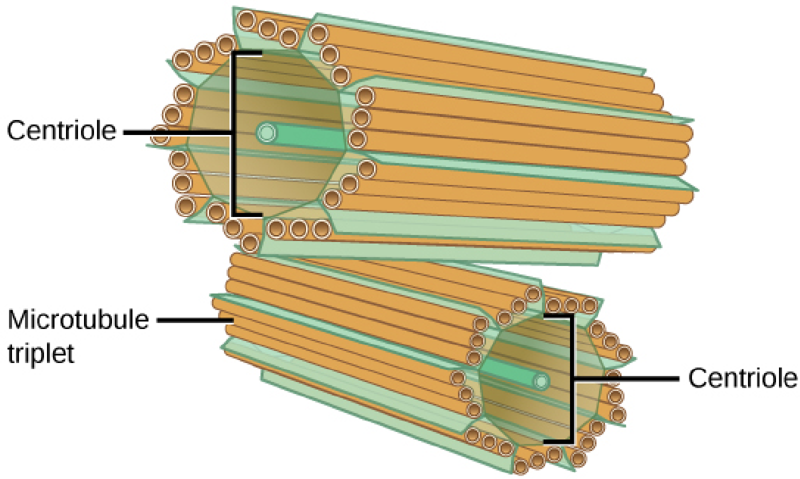
figure indicate a series of many more glucose units. The size of the page makes it impossible to portray an entire cellulose molecule. The chloroplast has an outer membrane, an inner membrane, and membrane structures - thylakoids that are stacked into grana. We call the space inside the thylakoid membranes the thylakoid space. The light harvesting reactions take place in the thylakoid membranes, and sugar synthesis takes place in the fluid inside the inner membrane, which we call the stroma. Chloroplasts also have their own genome, which is contained on a single circular chromosome.

Animal Cells versus Plant Cells

At this point, you know that each eukaryotic cell has a plasma membrane, cytoplasm, a nucleus, ribosomes, mitochondria, peroxisomes, and in some, vacuoles, but there are some striking differences between animal and plant cells. While both animal and plant cells have microtubule organizing centers (MTOCs), animal cells also have centrioles associated with the MTOC: a complex we call the centrosome. Animal cells each have a centrosome and lysosomes; whereas, most plant cells do not. Plant cells have a cell wall, chloroplasts and other specialized plastids, and a large central vacuole; whereas, animal cells do not.

The Centrosome

The **centrosome** is a microtubule-organizing center found near the nuclei of animal cells. It contains a pair of centrioles, two structures that lie perpendicular to each other ([link]). Each centriole is a cylinder of nine triplets of microtubules.



The centrosome (the organelle where all microtubules originate) replicates itself before a cell divides, and the centrioles appear to have some role in pulling the duplicated chromosomes to opposite ends of the dividing cell. However, the centriole's exact function in cell division isn’t clear, because cells that have had the centrosome removed can still divide, and plant cells, which lack centrosomes, are capable of cell division.

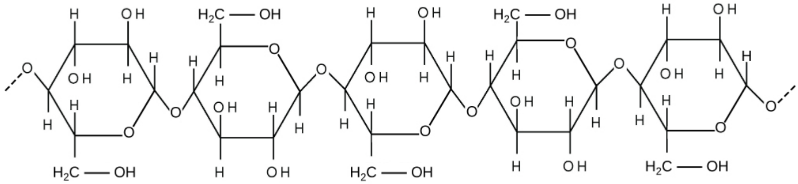
Lysosomes

Animal cells have another set of organelles that most plant cells do not: lysosomes. The **lysosomes**

are the cell’s “garbage disposal.” In plant cells, the digestive processes take place in vacuoles. Enzymes within the lysosomes aid in breaking down proteins, polysaccharides, lipids, nucleic acids, and even worn-out organelles. These enzymes are active at a much lower pH than the cytoplasm's. Therefore, the pH within lysosomes is more acidic than the cytoplasm's pH. Many reactions that take place in the cytoplasm could not occur at a low pH, so again, the advantage of compartmentalizing the eukaryotic cell into organelles is apparent.

The Cell Wall

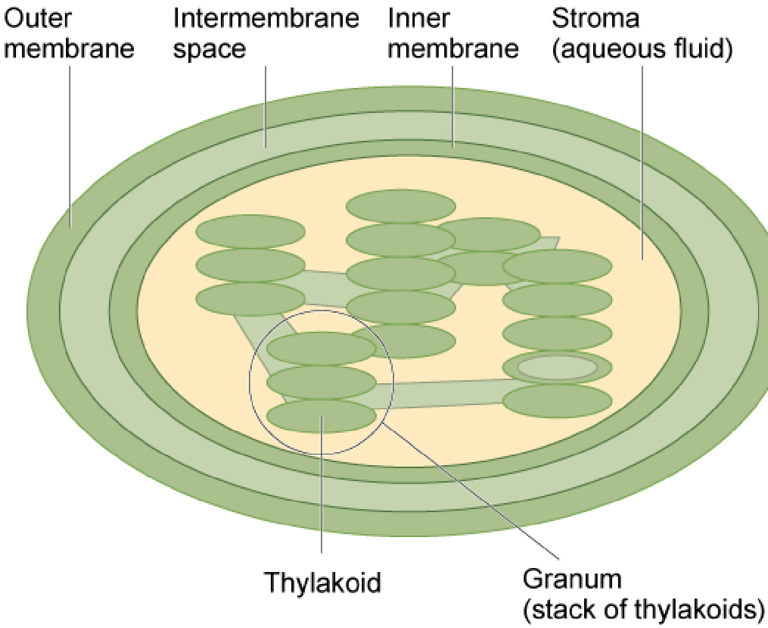
If you examine [link], the plant cell diagram, you will see a structure external to the plasma membrane. This is the **cell wall**, a rigid covering that protects the cell, provides structural support, and gives shape to the cell. Fungal and some protistan cells also have cell walls. While the prokaryotic cell walls' chief component is peptidoglycan, the major organic molecule in the plant (and some protists') cell wall is cellulose ([link]), a polysaccharide comprised of glucose units. Have you ever noticed that when you bite into a raw vegetable, like celery, it crunches? That’s because you are tearing the celery cells' rigid cell walls with your teeth.



Chloroplasts

Like the mitochondria, chloroplasts have their own DNA and ribosomes, but chloroplasts have an entirely different function. **Chloroplasts** are plant cell organelles that carry out photosynthesis. Photosynthesis is the series of reactions that use carbon dioxide, water, and light energy to make glucose and oxygen. This is a major difference between plants and animals. Plants (autotrophs) are able to make their own food, like sugars, while animals (heterotrophs) must ingest their food.

Like mitochondria, chloroplasts have outer and inner membranes, but within the space enclosed by a chloroplast's inner membrane is a set of interconnected and stacked fluid-filled membrane sacs we call thylakoids ([link]). Each thylakoid stack is a granum (plural = grana). We call the fluid enclosed by the inner membrane that surrounds the grana the stroma.



The chloroplasts contain a green pigment, **chlorophyll**, which captures the light energy that drives the reactions of photosynthesis. Like plant cells, photosynthetic protists also have chloroplasts. Some bacteria perform photosynthesis, but their chlorophyll is not relegated to an organelle.

Evolution Connection

Endosymbiosis

We have mentioned that both mitochondria and chloroplasts contain DNA and ribosomes. Have you wondered why? Strong evidence points to endosymbiosis as the explanation.

Symbiosis is a relationship in which organisms

from two separate species depend on each other for their survival. Endosymbiosis (endo- = “within”) is a mutually beneficial relationship in which one organism lives inside the other. Endosymbiotic relationships abound in nature. We have already mentioned that microbes that produce vitamin K live inside the human gut. This relationship is beneficial for us because we are unable to synthesize vitamin K. It is also beneficial for the microbes because they are protected from other organisms and from drying out, and they receive abundant food from the environment of the large intestine.

Scientists have long noticed that bacteria, mitochondria, and chloroplasts are similar in size. We also know that bacteria have DNA and ribosomes, just like mitochondria and chloroplasts. Scientists believe that host cells and bacteria formed an endosymbiotic relationship when the host cells ingested both aerobic and autotrophic bacteria (cyanobacteria) but did not destroy them. Through many millions of years of evolution, these ingested bacteria became more specialized in their functions, with the aerobic bacteria becoming mitochondria and the autotrophic bacteria becoming chloroplasts.

The Central Vacuole

Previously, we mentioned vacuoles as essential components of plant cells. If you look at [link]**b**, you will see that plant cells each have a large central vacuole that occupies most of the cell's area. The **central vacuole** plays a key role in regulating the cell's concentration of water in changing environmental conditions. Have you ever noticed that if you forget to water a plant for a few days, it wilts? That's because as the water concentration in the soil becomes lower than the water concentration in the plant, water moves out of the central vacuoles and cytoplasm. As the central vacuole shrinks, it leaves the cell wall unsupported. This loss of support to the plant's cell walls results in the wilted appearance.

The central vacuole also supports the cell's expansion. When the central vacuole holds more water, the cell becomes larger without having to invest considerable energy in synthesizing new cytoplasm.

Section Summary

Like a prokaryotic cell, a eukaryotic cell has a plasma membrane, cytoplasm, and ribosomes, but a eukaryotic cell is typically larger than a prokaryotic cell, has a true nucleus (meaning a membrane surrounds its DNA), and has other membrane-bound organelles that allow for compartmentalizing

functions. The plasma membrane is a phospholipid bilayer embedded with proteins. The nucleus's nucleolus is the site of ribosome assembly. We find ribosomes either in the cytoplasm or attached to the cytoplasmic side of the plasma membrane or endoplasmic reticulum. They perform protein synthesis. Mitochondria participate in cellular respiration. They are responsible for the majority of ATP produced in the cell. Peroxisomes hydrolyze fatty acids, amino acids, and some toxins. Vesicles and vacuoles are storage and transport compartments. In plant cells, vacuoles also help break down macromolecules.

Animal cells also have a centrosome and lysosomes. The centrosome has two bodies perpendicular to each other, the centrioles, and has an unknown purpose in cell division. Lysosomes are the digestive organelles of animal cells.

Plant cells and plant-like cells each have a cell wall, chloroplasts, and a central vacuole. The plant cell wall, whose primary component is cellulose, protects the cell, provides structural support, and gives the cell shape. Photosynthesis takes place in chloroplasts. The central vacuole can expand without having to produce more cytoplasm.

Visual Connection Questions

[link] If the nucleolus were not able to carry out its function, what other cellular organelles would be affected?

[link] Free ribosomes and rough endoplasmic reticulum (which contains ribosomes) would not be able to form.

Review Questions

Which of the following is surrounded by two phospholipid bilayers?

1. the ribosomes
2. the vesicles
3. the cytoplasm
4. the nucleoplasm

D

Peroxisomes got their name because hydrogen peroxide is:

1. used in their detoxification reactions
2. produced during their oxidation reactions

3. incorporated into their membranes
4. a cofactor for the organelles' enzymes

B

In plant cells, the function of the lysosomes is carried out by _______.

1. vacuoles
2. peroxisomes
3. ribosomes
4. nuclei

A

Which of the following is both in eukaryotic and prokaryotic cells?

1. nucleus
2. mitochondrion
3. vacuole
4. ribosomes

D

Tay-Sachs disease is a genetic disorder that

results in the destruction of neurons due to a buildup of sphingolipids in the cells. Which organelle is malfunctioning in Tay-Sachs?

1. lysosome
2. endoplasmic reticulum
3. peroxisome
4. mitochondria

A

Critical Thinking Questions

You already know that ribosomes are abundant in red blood cells. In what other cells of the body would you find them in great abundance? Why?

Ribosomes are abundant in muscle cells as well because muscle cells are constructed of the proteins made by the ribosomes.

What are the structural and functional similarities and differences between mitochondria and chloroplasts?

Both are similar in that they are enveloped in a double membrane, both have an intermembrane space, and both make ATP. Both mitochondria and chloroplasts have DNA, and mitochondria have inner folds called cristae and a matrix, while chloroplasts have chlorophyll and accessory pigments in the thylakoids that form stacks (grana) and a stroma.

Why are plasma membranes arranged as a bilayer rather than a monolayer?

The plasma membrane is a bilayer because the phospholipids that create it are amphiphilic (hydrophilic head, hydrophobic tail). If the plasma membrane was a monolayer, the hydrophobic tails of the phospholipids would be in direct contact with the inside of the cell. Since the cytoplasm is largely made of water, this interaction would not be stable, and would disrupt the plasma membrane of the cell as the tails were repulsed by the cytoplasm (in water, phospholipids spontaneously form spherical droplets with the hydrophilic heads facing outward to isolate the hydrophobic tails from the water). By having a bilayer, the hydrophilic heads are exposed to the aqueous cytoplasm and extracellular space, while the hydrophobic

tails interact with each other in the middle of the membrane.

Glossary

cell wall
: rigid cell covering comprised of various molecules that protects the cell, provides structural support, and gives shape to the cell

central vacuole
: large plant cell organelle that regulates the cell's storage compartment, holds water, and plays a significant role in cell growth as the site of macromolecule degradation

centrosome
: region in animal cells made of two centrioles that serves as an organizing center for microtubules

chlorophyll
: green pigment that captures the light energy that drives the light reactions of photosynthesis

chloroplast
: plant cell organelle that carries out photosynthesis

chromatin

protein-DNA complex that serves as the chromosomes' building material

chromosome
: structure within the nucleus that comprises chromatin that contains DNA, the hereditary material

cytoplasm
: entire region between the plasma membrane and the nuclear envelope, consisting of organelles suspended in the gel-like cytosol, the cytoskeleton, and various chemicals

cytosol
: the cytoplasm's gel-like material in which cell structures are suspended

eukaryotic cell
: cell that has a membrane-bound nucleus and several other membrane-bound compartments or sacs

lysosome
: organelle in an animal cell that functions as the cell’s digestive component; it breaks down proteins, polysaccharides, lipids, nucleic acids, and even worn-out organelles

mitochondria
: (singular = mitochondrion) cellular organelles responsible for carrying out

cellular respiration, resulting in producing ATP, the cell’s main energy-carrying molecule

nuclear envelope
: double-membrane structure that constitutes the nucleus' outermost portion

nucleolus
: darkly staining body within the nucleus that is responsible for assembling ribosome subunits

nucleoplasm
: semi-solid fluid inside the nucleus that contains the chromatin and nucleolus

nucleus
: cell organelle that houses the cell’s DNA and directs ribosome and protein synthesis

organelle
: compartment or sac within a cell

peroxisome
: small, round organelle that contains hydrogen peroxide, oxidizes fatty acids and amino acids, and detoxifies many poisons

plasma membrane
: phospholipid bilayer with embedded (integral) or attached (peripheral) proteins, and separates the cell's internal content from

its surrounding environment

ribosome
: cellular structure that carries out protein synthesis

vacuole
: membrane-bound sac, somewhat larger than a vesicle, which functions in cellular storage and transport

vesicle
: small, membrane-bound sac that functions in cellular storage and transport; its membrane is capable of fusing with the plasma membrane and the membranes of the endoplasmic reticulum and Golgi apparatus

The Endomembrane System and Proteins

By the end of this section, you will be able to do the following:

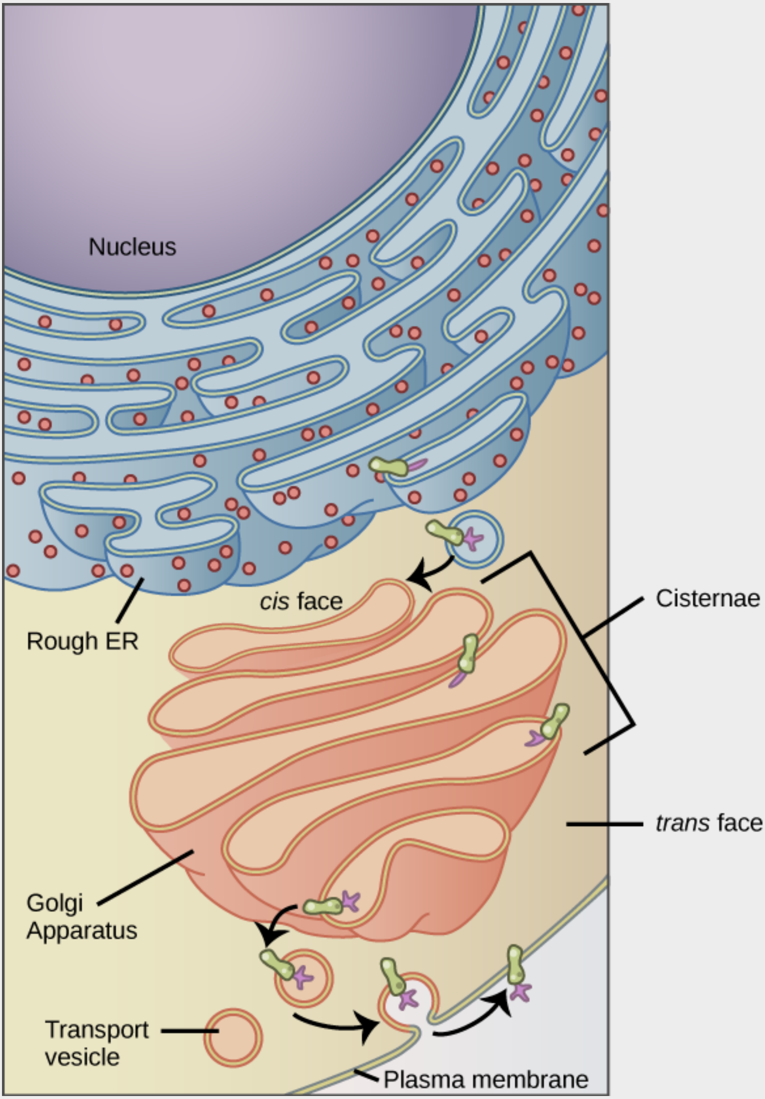
- List the components of the endomembrane system
- Recognize the relationship between the endomembrane system and its functions

The endomembrane system (endo = "within") is a group of membranes and organelles ([link]) in eukaryotic cells that works together to modify, package, and transport lipids and proteins. It includes the nuclear envelope, lysosomes, and vesicles, which we have already mentioned, and the endoplasmic reticulum and Golgi apparatus, which we will cover shortly. Although not technically *within* the cell, the plasma membrane is included in the endomembrane system because, as you will see, it interacts with the other endomembranous organelles. The endomembrane system does not include either mitochondria or chloroplast membranes.

Visual Connection

Membrane and secretory proteins are synthesized in the rough endoplasmic reticulum (RER). The RER also sometimes modifies proteins. In this illustration, an attachment of a (purple)

carbohydrate modifies a (green) integral membrane protein in the ER. Vesicles with the integral protein bud from the ER and fuse with the Golgi apparatus' cis face. As the protein passes along the Golgi's cisternae, the addition of more carbohydrates further modifies it. After its synthesis is complete, it exits as an integral membrane protein of the vesicle that buds from the Golgi's **trans** face. When the vesicle fuses with the cell membrane, the protein becomes an integral portion of that cell membrane. (credit: modification of work by Magnus Manske)



If a peripheral membrane protein were synthesized in the lumen (inside) of the ER, would it end up on the inside or outside of the plasma membrane?

This transmission electron micrograph shows the

rough endoplasmic reticulum and other organelles in a pancreatic cell. (credit: modification of work by Louisa Howard)

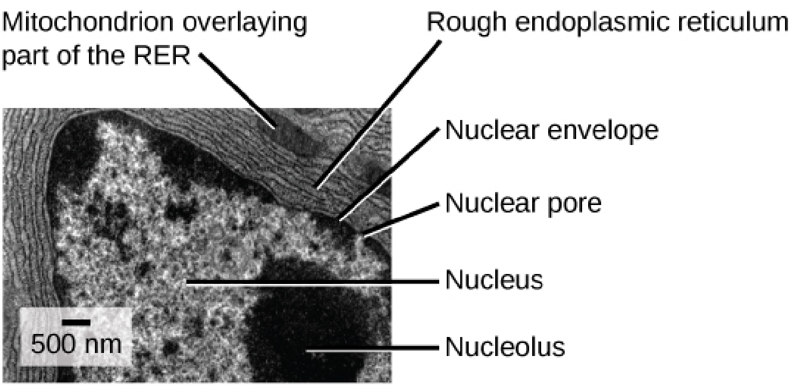
The Endoplasmic Reticulum

The **endoplasmic reticulum (ER)** ([link]) is a series of interconnected membranous sacs and tubules that collectively modifies proteins and synthesizes lipids. However, these two functions take place in separate areas of the ER: the rough ER and the smooth ER, respectively.

We call the ER tubules' hollow portion the lumen or cisternal space. The ER's membrane, which is a phospholipid bilayer embedded with proteins, is continuous with the nuclear envelope.

Rough ER

Scientists have named the **rough endoplasmic reticulum (RER)** as such because the ribosomes attached to its cytoplasmic surface give it a studded appearance when viewing it through an electron microscope ([link]).



Ribosomes transfer their newly synthesized proteins into the RER's lumen where they undergo structural modifications, such as folding or acquiring side chains. These modified proteins incorporate into cellular membranes—the ER or the ER's or other organelles' membranes. The proteins can also secrete from the cell (such as protein hormones, enzymes). The RER also makes phospholipids for cellular membranes.

If the phospholipids or modified proteins are not destined to stay in the RER, they will reach their destinations via transport vesicles that bud from the RER's membrane ([link]).

Since the RER is engaged in modifying proteins (such as enzymes, for example) that secrete from the cell, you would be correct in assuming that the RER is abundant in cells that secrete proteins. This is the case with liver cells, for example.

Smooth ER

The **smooth endoplasmic reticulum (SER)** is continuous with the RER but has few or no ribosomes on its cytoplasmic surface ([link]). SER functions include synthesis of carbohydrates, lipids, and steroid hormones; detoxification of medications and poisons; and storing calcium ions.

In muscle cells, a specialized SER, the sarcoplasmic reticulum, is responsible for storing calcium ions that are needed to trigger the muscle cells' coordinated contractions.

Link to Learning

You can watch an excellent animation of the endomembrane system here. At the end of the animation, there is a short self-assessment.

Career Connection

Cardiologist

Heart disease is the leading cause of death in the United States. This is primarily due to our sedentary lifestyle and our high trans-fat diets. Heart failure is just one of many disabling heart conditions. Heart failure does not mean that the heart has stopped working. Rather, it means that the heart can’t pump with sufficient force to transport oxygenated blood to all the vital organs.

Left untreated, heart failure can lead to kidney failure and other organ failure.

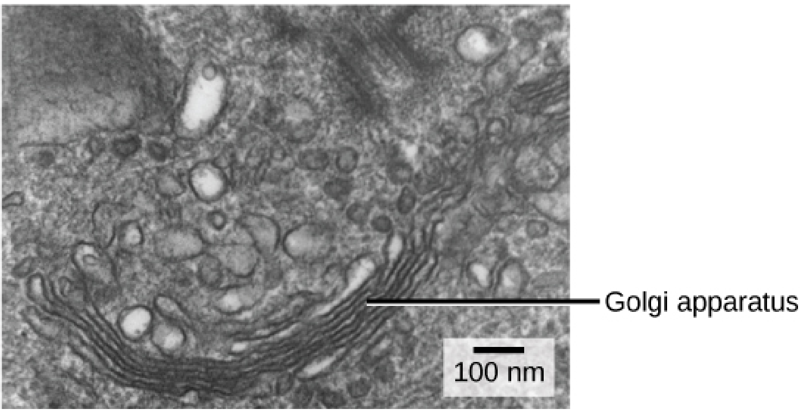
Cardiac muscle tissue comprises the heart's wall. Heart failure occurs when cardiac muscle cells' endoplasmic reticula do not function properly. As a result, an insufficient number of calcium ions are available to trigger a sufficient contractile force. Cardiologists (cardi- = “heart”; -ologist = “one who studies”) are doctors who specialize in treating heart diseases, including heart failure. Cardiologists can diagnose heart failure via a physical examination, results from an electrocardiogram (ECG, a test that measures the heart's electrical activity), a chest X-ray to see whether the heart is enlarged, and other tests. If the cardiologist diagnoses heart failure, he or she will typically prescribe appropriate medications and recommend a reduced table salt intake and a supervised exercise program.

The Golgi apparatus in this white blood cell is visible as a stack of semicircular, flattened rings in the lower portion of the image. You can see several vesicles near the Golgi apparatus. (credit: modification of work by Louisa Howard)

The Golgi Apparatus

We have already mentioned that vesicles can bud

from the ER and transport their contents elsewhere, but where do the vesicles go? Before reaching their final destination, the lipids or proteins within the transport vesicles still need sorting, packaging, and tagging so that they end up in the right place. Sorting, tagging, packaging, and distributing lipids and proteins takes place in the **Golgi apparatus** (also called the Golgi body), a series of flattened membranes ([link]).



We call the Golgi apparatus' the *cis* face. The opposite side is the *trans* face. The transport vesicles that formed from the ER travel to the *cis* face, fuse with it, and empty their contents into the Golgi apparatus' lumen. As the proteins and lipids travel through the Golgi, they undergo further modifications that allow them to be sorted. The most frequent modification is adding short sugar molecule chains. These newly modified proteins and lipids then tag with phosphate groups or other small molecules in order to travel to their proper destinations.

Finally, the modified and tagged proteins are packaged into secretory vesicles that bud from the Golgi's *trans* face. While some of these vesicles deposit their contents into other cell parts where they will be used, other secretory vesicles fuse with the plasma membrane and release their contents outside the cell.

In another example of form following function, cells that engage in a great deal of secretory activity (such as salivary gland cells that secrete digestive enzymes or immune system cells that secrete antibodies) have an abundance of Golgi.

In plant cells, the Golgi apparatus has the additional role of synthesizing polysaccharides, some of which are incorporated into the cell wall and some of which other cell parts use.

Career Connection

Geneticist

Many diseases arise from genetic mutations that prevent synthesizing critical proteins. One such disease is Lowe disease (or oculocerebrorenal syndrome, because it affects the eyes, brain, and kidneys). In Lowe disease, there is a deficiency in an enzyme localized to the Golgi apparatus. Children with Lowe disease are born with cataracts, typically develop kidney disease after the

first year of life, and may have impaired mental abilities.

A mutation on the X chromosome causes Lowe disease. The X chromosome is one of the two human sex chromosomes, as these chromosomes determine a person's sex. Females possess two X chromosomes while males possess one X and one Y chromosome. In females, the genes on only one of the two X chromosomes are expressed. Females who carry the Lowe disease gene on one of their X chromosomes are carriers and do not show symptoms of the disease. However, males only have one X chromosome and the genes on this chromosome are always expressed. Therefore, males will always have Lowe disease if their X chromosome carries the Lowe disease gene. Geneticists have identified the mutated gene's location, as well as many other mutation locations that cause genetic diseases. Through prenatal testing, a woman can find out if the fetus she is carrying may be afflicted with one of several genetic diseases.

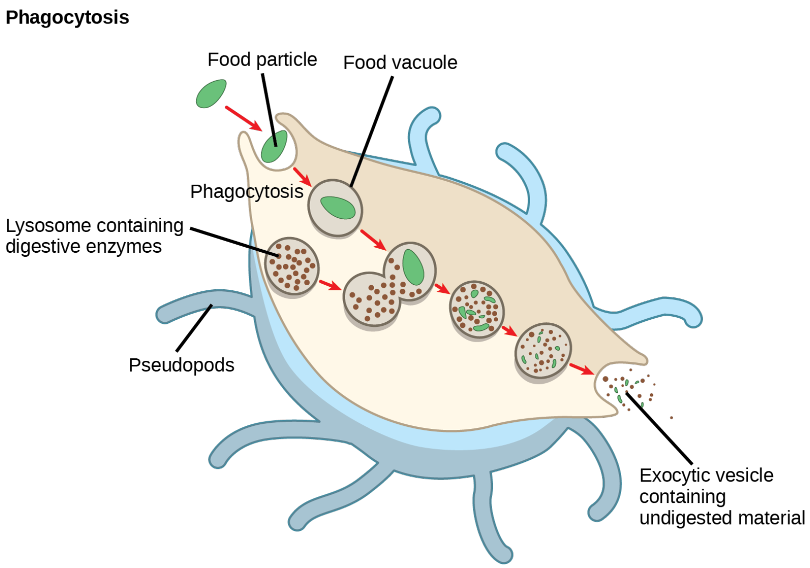
Geneticists analyze prenatal genetic test results and may counsel pregnant women on available options. They may also conduct genetic research that leads to new drugs or foods, or perform DNA analyses for forensic investigations.

A macrophage has engulfed (phagocytized) a

potentially pathogenic bacterium and then fuses with lysosomes within the cell to destroy the pathogen. Other organelles are present in the cell but for simplicity we do not show them.

Lysosomes

In addition to their role as the digestive component and organelle-recycling facility of animal cells, lysosomes are part of the endomembrane system. Lysosomes also use their hydrolytic enzymes to destroy pathogens (disease-causing organisms) that might enter the cell. A good example of this occurs in macrophages, a group of white blood cells which are part of your body’s immune system. In a process that scientists call phagocytosis or endocytosis, a section of the macrophage's plasma membrane invaginates (folds in) and engulfs a pathogen. The invaginated section, with the pathogen inside, then pinches itself off from the plasma membrane and becomes a vesicle. The vesicle fuses with a lysosome. The lysosome’s hydrolytic enzymes then destroy the pathogen ([link]).



Section Summary

The endomembrane system includes the nuclear envelope, lysosomes, vesicles, the ER, and Golgi apparatus, as well as the plasma membrane. These cellular components work together to modify, package, tag, and transport proteins and lipids that form the membranes.

The RER modifies proteins and synthesizes phospholipids in cell membranes. The SER synthesizes carbohydrates, lipids, and steroid hormones; engages in the detoxification of medications and poisons; and stores calcium ions. Sorting, tagging, packaging, and distributing lipids and proteins take place in the Golgi apparatus.

Budding RER and Golgi membranes create lysosomes. Lysosomes digest macromolecules, recycle worn-out organelles, and destroy pathogens.

Visual Connection Questions

[link] If a peripheral membrane protein were synthesized in the lumen (inside) of the ER, would it end up on the inside or outside of the plasma membrane?

[link] It would end up on the outside. After the vesicle passes through the Golgi apparatus and fuses with the plasma membrane, it turns inside out.

Review Questions

Which of the following is not a component of the endomembrane system?

1. mitochondrion
2. Golgi apparatus
3. endoplasmic reticulum

4. lysosome

A

The process by which a cell engulfs a foreign particle is known as:

1. endosymbiosis
2. phagocytosis
3. hydrolysis
4. membrane synthesis

B

Which of the following is most likely to have the greatest concentration of smooth endoplasmic reticulum?

1. a cell that secretes enzymes
2. a cell that destroys pathogens
3. a cell that makes steroid hormones
4. a cell that engages in photosynthesis

C

Which of the following sequences correctly lists

in order the steps involved in the incorporation of a proteinaceous molecule within a cell?

1. protein synthesis of the protein on the ribosome; modification in the Golgi apparatus; packaging in the endoplasmic reticulum; tagging in the vesicle
2. synthesis of the protein on the lysosome; tagging in the Golgi; packaging in the vesicle; distribution in the endoplasmic reticulum
3. synthesis of the protein on the ribosome; modification in the endoplasmic reticulum; tagging in the Golgi; distribution via the vesicle
4. synthesis of the protein on the lysosome; packaging in the vesicle; distribution via the Golgi; tagging in the endoplasmic reticulum

---

C

Congenital disorders of glycosylation are a growing class of rare diseases. Which organelle would be most commonly involved in the glycoprotein disorder portion of the group?

1. RER
2. ribosomes
3. endosomes

4. Golgi apparatus

D

Critical Thinking Questions

In the context of cell biology, what do we mean by form follows function? What are at least two examples of this concept?

"Form follows function" refers to the idea that the function of a body part dictates the form of that body part. As an example, compare your arm to a bat's wing. While the bones of the two correspond, the parts serve different functions in each organism and their forms have adapted to follow that function.

In your opinion, is the nuclear membrane part of the endomembrane system? Why or why not? Defend your answer.

Since the external surface of the nuclear membrane is continuous with the rough endoplasmic reticulum, which is part of the

endomembrane system, then it is correct to say that it is part of the system.

Glossary

endomembrane system
: group of organelles and membranes in eukaryotic cells that work together modifying, packaging, and transporting lipids and proteins

endoplasmic reticulum (ER)
: series of interconnected membranous structures within eukaryotic cells that collectively modify proteins and synthesize lipids

Golgi apparatus
: eukaryotic organelle comprised of a series of stacked membranes that sorts, tags, and packages lipids and proteins for distribution

rough endoplasmic reticulum (RER)
: region of the endoplasmic reticulum that is studded with ribosomes and engages in protein modification and phospholipid synthesis

smooth endoplasmic reticulum (SER)
: region of the endoplasmic reticulum that has few or no ribosomes on its cytoplasmic

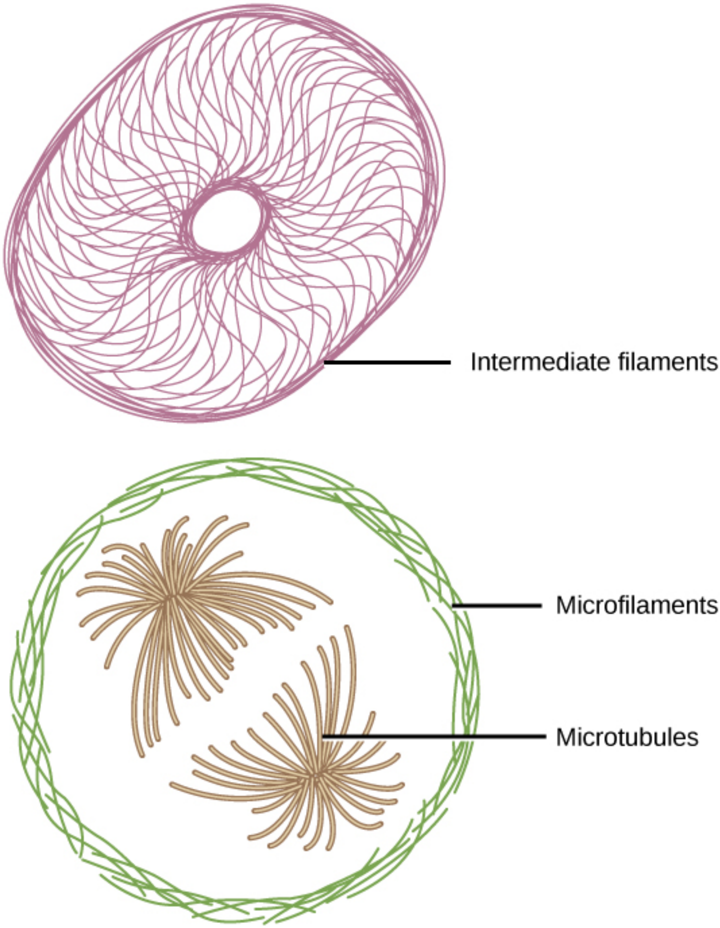
surface and synthesizes carbohydrates, lipids, and steroid hormones; detoxifies certain chemicals (like pesticides, preservatives, medications, and environmental pollutants), and stores calcium ions

The Cytoskeleton

By the end of this section, you will be able to do the following:

- Describe the cytoskeleton
- Compare the roles of microfilaments, intermediate filaments, and microtubules
- Compare and contrast cilia and flagella
- Summarize the differences among the components of prokaryotic cells, animal cells, and plant cells

If you were to remove all the organelles from a cell, would the plasma membrane and the cytoplasm be the only components left? No. Within the cytoplasm, there would still be ions and organic molecules, plus a network of protein fibers that help maintain the cell's shape, secure some organelles in specific positions, allow cytoplasm and vesicles to move within the cell, and enable cells within multicellular organisms to move. Collectively, scientists call this network of protein fibers the **cytoskeleton**. There are three types of fibers within the cytoskeleton: microfilaments, intermediate filaments, and microtubules ([link]). Here, we will examine each. Microfilaments thicken the cortex around the cell's inner edge. Like rubber bands, they resist tension. There are microtubules in the cell's interior where they maintain their shape by resisting compressive forces. There are intermediate filaments throughout the cell that hold organelles in place.

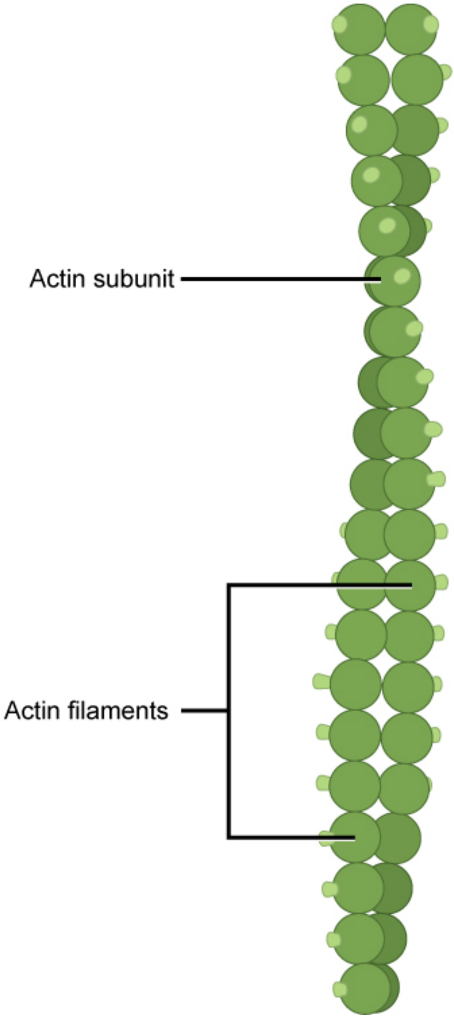


Two intertwined actin strands comprise microfilaments.

Microfilaments

Of the three types of protein fibers in the cytoskeleton, **microfilaments** are the narrowest. They function in cellular movement, have a diameter of about 7 nm, and are comprised of two globular protein intertwined strands, which we call

actin ([link]). For this reason, we also call microfilaments actin filaments.



ATP powers actin to assemble its filamentous form, which serves as a track for the movement of a motor protein we call myosin. This enables actin to engage in cellular events requiring motion, such as cell division in eukaryotic cells and cytoplasmic streaming, which is the cell cytoplasm's circular

movement in plant cells. Actin and myosin are plentiful in muscle cells. When your actin and myosin filaments slide past each other, your muscles contract.

Microfilaments also provide some rigidity and shape to the cell. They can depolymerize (disassemble) and reform quickly, thus enabling a cell to change its shape and move. White blood cells (your body's infection-fighting cells) make good use of this ability. They can move to an infection site and phagocytize the pathogen.

Link to Learning

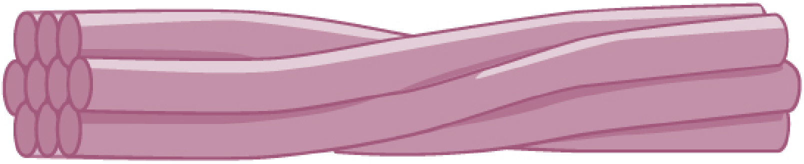
To see an example of a white blood cell in action, watch a short time-lapse video of the cell capturing two bacteria. It engulfs one and then moves on to the other.
https://www.openstax.org/l/chasing_bcteria

Intermediate filaments consist of several intertwined strands of fibrous proteins.

Intermediate Filaments

Several strands of fibrous proteins that are wound together comprise intermediate filaments ([link]).

Cytoskeleton elements get their name from the fact that their diameter, 8 to 10 nm, is between those of microfilaments and microtubules.



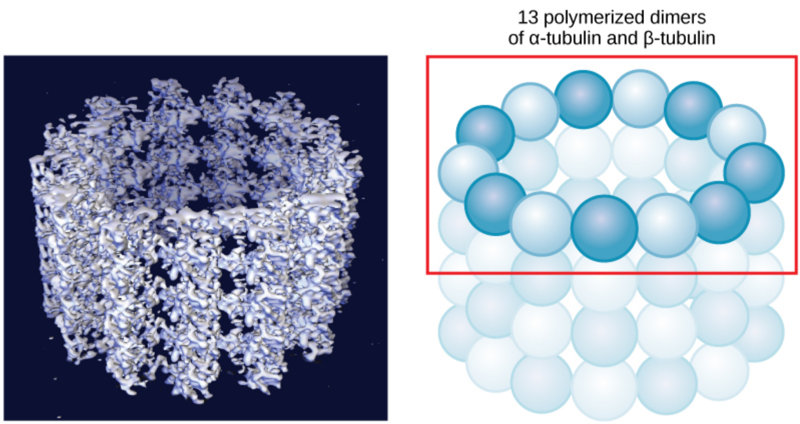
Intermediate filaments have no role in cell movement. Their function is purely structural. They bear tension, thus maintaining the cell's shape, and anchor the nucleus and other organelles in place. [link] shows how intermediate filaments create a supportive scaffolding inside the cell.

The intermediate filaments are the most diverse group of cytoskeletal elements. Several fibrous protein types are in the intermediate filaments. You are probably most familiar with keratin, the fibrous protein that strengthens your hair, nails, and the skin's epidermis.

Microtubules are hollow. Their walls consist of 13 polymerized dimers of α-tubulin and β-tubulin (right image). The left image shows the tube's molecular structure. This transmission electron micrograph of two flagella shows the microtubules' 9 + 2 array: nine microtubule doublets surround a single microtubule doublet. (credit: modification of work by Dartmouth Electron Microscope Facility, Dartmouth College; scale-bar data from Matt Russell)

Microtubules

As their name implies, microtubules are small hollow tubes. Polymerized dimers of α-tubulin and β-tubulin, two globular proteins, comprise the microtubule's walls ([link]). With a diameter of about 25 nm, **microtubules** are cytoskeletons' widest components. They help the cell resist compression, provide a track along which vesicles move through the cell, and pull replicated chromosomes to opposite ends of a dividing cell. Like microfilaments, microtubules can disassemble and reform quickly.

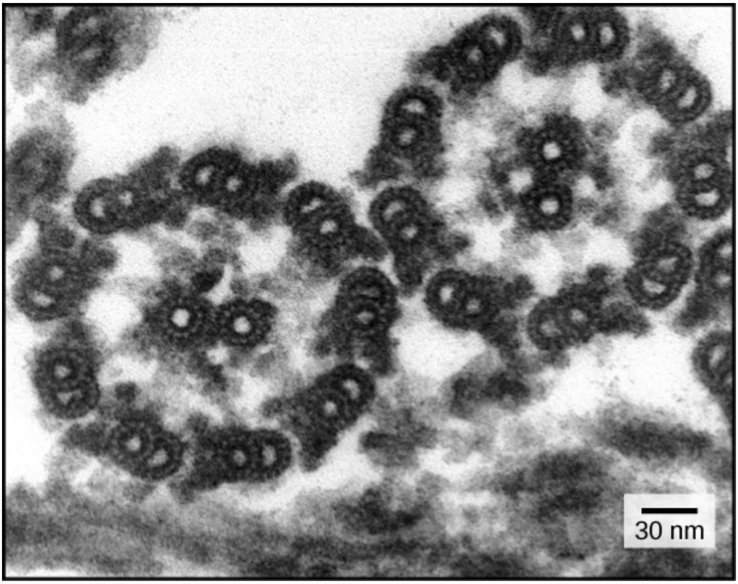


Microtubules are also the structural elements of flagella, cilia, and centrioles (the latter are the centrosome's two perpendicular bodies). In animal cells, the centrosome is the microtubule-organizing center. In eukaryotic cells, flagella and cilia are quite different structurally from their counterparts in prokaryotes, as we discuss below.

Flagella and Cilia

The **flagella** (singular = flagellum) are long, hair-like structures that extend from the plasma membrane and enable an entire cell to move (for example, sperm, *Euglena*, and some prokaryotes). When present, the cell has just one flagellum or a few flagella. However, when **cilia** (singular = cilium) are present, many of them extend along the plasma membrane's entire surface. They are short, hair-like structures that move entire cells (such as paramecia) or substances along the cell's outer surface (for example, the cilia of cells lining the Fallopian tubes that move the ovum toward the uterus, or cilia lining the cells of the respiratory tract that trap particulate matter and move it toward your nostrils.)

Despite their differences in length and number, flagella and cilia share a common structural arrangement of microtubules called a “9 + 2 array.” This is an appropriate name because a single flagellum or cilium is made of a ring of nine microtubule doublets, surrounding a single microtubule doublet in the center ([link]).



You have now completed a broad survey of prokaryotic and eukaryotic cell components. For a summary of cellular components in prokaryotic and eukaryotic cells, see [link].

Components of Prokaryotic and Eukaryotic Cells

Cell Component	Function	Present in	Present in	Present in Plant

		Prokaryotes?	Animal Cells?	Cells?
Plasma membrane	Separates cell from external environment; controls passage of organic molecules, ions, water, oxygen, and wastes into and out of cell	Yes	Yes	Yes
Cytoplasm	Provides turgor pressure to plant cells as fluid inside the central vacuole; site of many metabolic reactions; medium	Yes	Yes	Yes

	in which organelles are found			
Nucleolus	Darkened area within the nucleus where ribosomal subunits are synthesized.	No	Yes	Yes
Nucleus	Cell organelle that houses DNA and directs synthesis of ribosomes and proteins	No	Yes	Yes
Ribosomes	Protein synthesis	Yes	Yes	Yes
Mitochondria	ATP production/ cellular respiration	No	Yes	Yes
Peroxisomes	Oxidize and thus	No	Yes	Yes

	break down fatty acids and amino acids, and detoxify poisons			
Vesicles and vacuoles	Storage and transport; digestive function in plant cells	No	Yes	Yes
Centrosome	Unspecified role in cell division in animal cells; microtubule source in animal cells	No	Yes	No
Lysosomes	Digestion of macromolecules; recycling of worn-out	No	Yes	Some

	organelles			
Cell wall	Protection, structural support, and maintenance of cell shape	Yes, primarily peptidoglycan	No	Yes, primarily cellulose
Chloroplasts	Photosynthesis	No	No	Yes
Endoplasmic reticulum	Modifies proteins and synthesizes lipids	No	Yes	Yes
Golgi apparatus	Modifies, sorts, tags, packages, and distributes lipids and proteins	No	Yes	Yes
Cytoskeleton	Maintains cell's shape, secures organelles in specific positions, allows cytoplasm	Yes	Yes	Yes

	and vesicles to move within cell, and enables unicellular organisms to move independently			
Flagella	Cellular locomotion	Some	Some	No, except for some plant sperm cells
Cilia	Cellular locomotion, movement of particles along plasma membrane's extracellular surface, and filtration	Some	Some	No

Section Summary

The cytoskeleton has three different protein element types. From narrowest to widest, they are the microfilaments (actin filaments), intermediate filaments, and microtubules. Biologists often associate microfilaments with myosin. They provide rigidity and shape to the cell and facilitate cellular movements. Intermediate filaments bear tension and anchor the nucleus and other organelles in place. Microtubules help the cell resist compression, serve as tracks for motor proteins that move vesicles through the cell, and pull replicated chromosomes to opposite ends of a dividing cell. They are also the structural element of centrioles, flagella, and cilia.

Review Questions

Which of the following have the ability to disassemble and reform quickly?

1. microfilaments and intermediate filaments
2. microfilaments and microtubules
3. intermediate filaments and microtubules
4. only intermediate filaments

B

Which of the following do not play a role in intracellular movement?

1. microfilaments and intermediate filaments
2. microfilaments and microtubules
3. intermediate filaments and microtubules
4. only intermediate filaments

D

In humans, ____ are used to move a cell within its environment while ____ are used to move the environment relative to the cell.

1. cilia, pseudopodia
2. flagella; cilia
3. microtubules; flagella
4. microfilaments; microtubules

B

Critical Thinking Questions

What are the similarities and differences between the structures of centrioles and

flagella?

Centrioles and flagella are alike in that they are made up of microtubules. In centrioles, two rings of nine microtubule “triplets” are arranged at right angles to one another. This arrangement does not occur in flagella.

How do cilia and flagella differ?

Cilia and flagella are alike in that they are made up of microtubules. Cilia are short, hair-like structures that exist in large numbers and usually cover the entire surface of the plasma membrane. Flagella, in contrast, are long, hair-like structures; when flagella are present, a cell has just one or two.

Describe how microfilaments and microtubules are involved in the phagocytosis and destruction of a pathogen by a macrophage.

A macrophage engulfs a pathogen by rearranging its actin microfilaments to bend the plasma membrane around the pathogen. Once the pathogen is sealed in an endosome inside the macrophage, the vesicle is walked along

microtubules until it combines with a lysosome to digest the pathogen.

Compare and contrast the boundaries that plant, animal, and bacteria cells use to separate themselves from their surrounding environment.

All three cell types have a plasma membrane that borders the cytoplasm on its interior side. In animal cells, the exterior side of the plasma membrane is in contact with the extracellular environment. However, in plant and bacteria cells, a cell wall surrounds the outside of the plasma membrane. In plants, the cell wall is made of cellulose, while in bacteria the cell wall is made of peptidoglycan. Gram-negative bacteria also have an additional capsule made of lipopolysaccharides that surrounds their cell wall.

Glossary

cilium
: (plural = cilia) short, hair-like structure that extends from the plasma membrane in large numbers and functions to move an entire cell or move substances along the cell's outer surface

cytoskeleton
: protein fiber network that collectively maintains the cell's shape, secures some organelles in specific positions, allows cytoplasm and vesicles to move within the cell, and enables unicellular organisms to move independently

flagellum
: (plural = flagella) long, hair-like structure that extends from the plasma membrane and moves the cell

intermediate filament
: cytoskeletal component, comprised of several fibrous protein intertwined strands, that bears tension, supports cell-cell junctions, and anchors cells to extracellular structures

microfilament
: the cytoskeleton system's narrowest element; it provides rigidity and shape to the cell and enables cellular movements

microtubule
: the cytoskeleton system's widest element; it helps the cell resist compression, provides a track along which vesicles move through the cell, pulls replicated chromosomes to opposite ends of a dividing cell, and is the structural element of centrioles, flagella, and cilia

Connections between Cells and Cellular Activities
By the end of this section, you will be able to do the following:

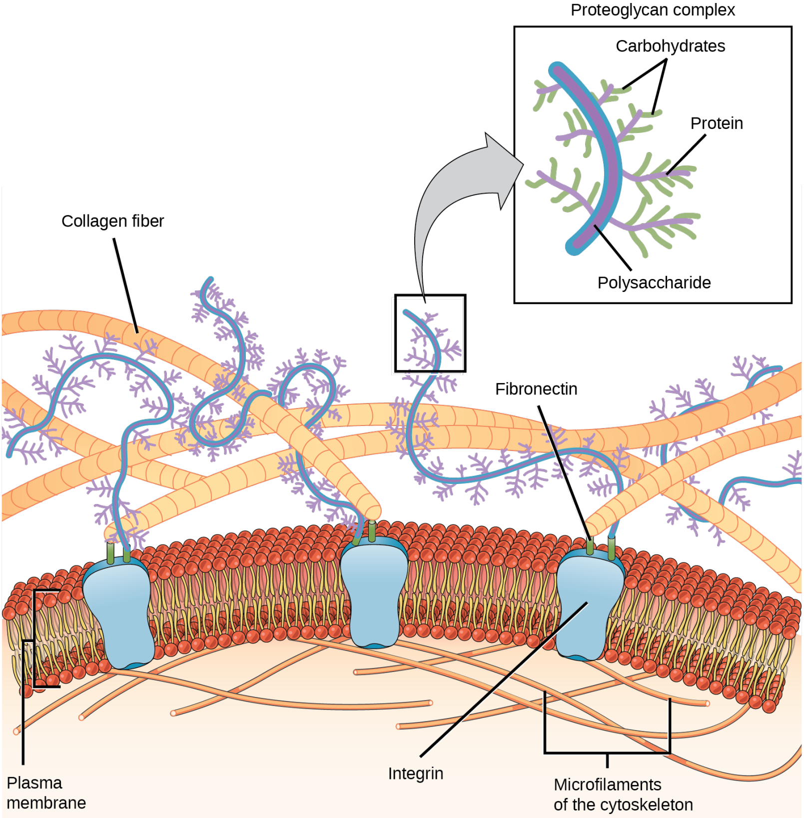
- Describe the extracellular matrix
- List examples of the ways that plant cells and animal cells communicate with adjacent cells
- Summarize the roles of tight junctions, desmosomes, gap junctions, and plasmodesmata

You already know that tissue is a group of similar cells working together. As you might expect, if cells are to work together, they must communicate with each other, just as you need to communicate with others if you work on a group project. Let's take a look at how cells communicate with each other.
The extracellular matrix consists of a network of proteins and carbohydrates.

Extracellular Matrix of Animal Cells

While cells in most multicellular organisms release materials into the extracellular space, animal cells will be discussed as an example. The primary components of these materials are proteins, and the most abundant protein is collagen. Collagen fibers are interwoven with proteoglycans, which are carbohydrate-containing protein molecules. Collectively, we call these materials the

extracellular matrix ([link]). Not only does the extracellular matrix hold the cells together to form a tissue, but it also allows the cells within the tissue to communicate with each other. How can this happen?



Cells have protein receptors on their plasma membranes' extracellular surfaces. When a molecule within the matrix binds to the receptor, it changes the receptor's molecular structure. The receptor, in turn, changes the microfilaments' conformation positioned just inside the plasma membrane. These conformational changes induce chemical signals

inside the cell that reach the nucleus and turn “on” or “off” the transcription of specific DNA sections, which affects the associated protein production, thus changing the activities within the cell.

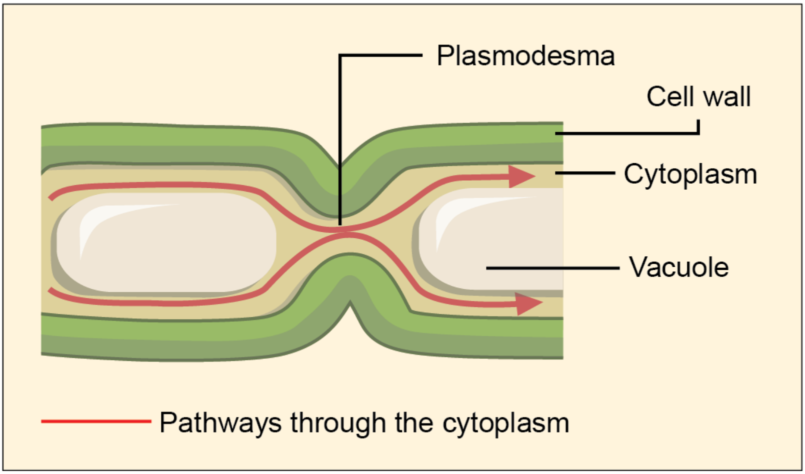
Blood clotting provides an example of the extracellular matrix's role in cell communication. When the cells lining a blood vessel are damaged, they display a protein receptor, which we call tissue factor. When tissue factor binds with another factor in the extracellular matrix, it causes platelets to adhere to the damaged blood vessel's wall, stimulates the adjacent smooth muscle cells in the blood vessel to contract (thus constricting the blood vessel), and initiates a series of steps that stimulate the platelets to produce clotting factors.
A plasmodesma is a channel between two adjacent plant cells' cell walls. Plasmodesmata allow materials to pass from one plant cell's cytoplasm to an adjacent cell's cytoplasm. Tight junctions form watertight connections between adjacent animal cells. Proteins create tight junction adherence. (credit: modification of work by Mariana Ruiz Villareal) A desmosome forms a very strong spot weld between cells. Linking cadherins and intermediate filaments create it. (credit: modification of work by Mariana Ruiz Villareal) A gap junction is a protein-lined pore that allows water and small molecules to pass between adjacent animal cells. (credit: modification of work by Mariana Ruiz Villareal)

Intercellular Junctions

Cells can also communicate with each other via direct contact, or intercellular junctions. There are differences in the ways that plant and animal and fungal cells communicate. Plasmodesmata are junctions between plant cells; whereas, animal cell contacts include tight junctions, gap junctions, and desmosomes.

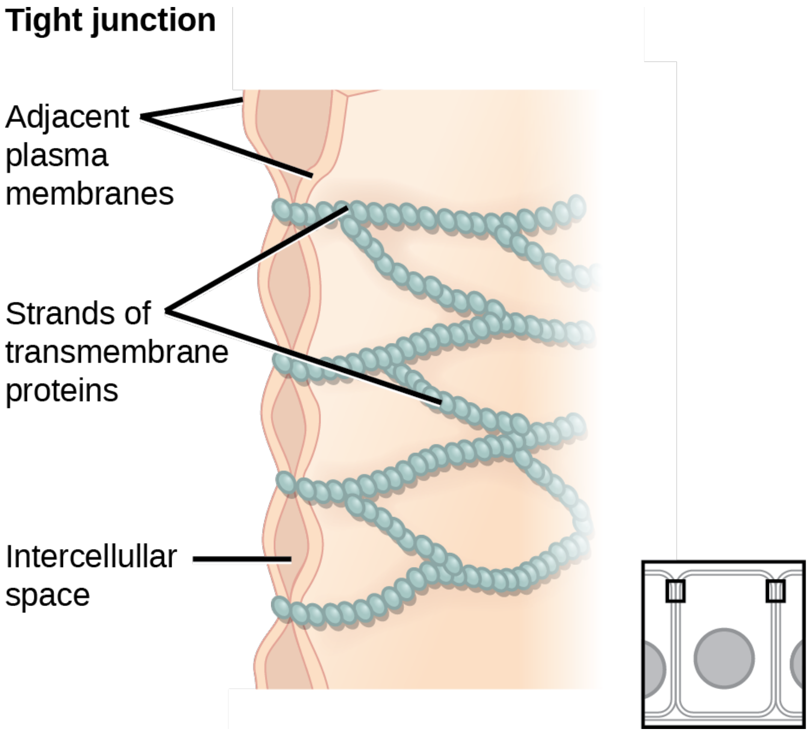
Plasmodesmata

In general, long stretches of the plasma membranes of neighboring plant cells cannot touch one another because the cell wall that surrounds each cell separates them ([link]). How then, can a plant transfer water and other soil nutrients from its roots, through its stems, and to its leaves? Such transport uses the vascular tissues (xylem and phloem) primarily. There also exist structural modifications, which we call **plasmodesmata** (singular = plasmodesma). Numerous channels that pass between adjacent plant cells' cell walls connect their cytoplasm, and enable transport of materials from cell to cell, and thus throughout the plant ([link]).



Tight Junctions

A **tight junction** is a watertight seal between two adjacent animal cells ([link]). Proteins (predominantly two proteins called claudins and occludins) tightly hold the cells against each other.

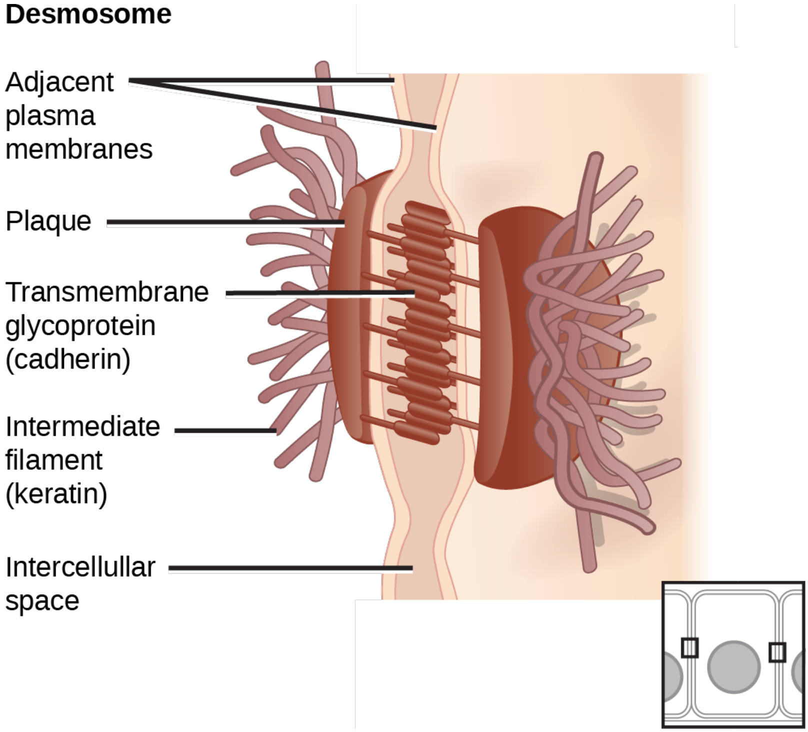


This tight adherence prevents materials from leaking between the cells; tight junctions are typically found in epithelial tissues that line internal organs and cavities, and comprise most of the skin. For example, the tight junctions of the epithelial cells lining your urinary bladder prevent urine from leaking out into the extracellular space.

Desmosomes

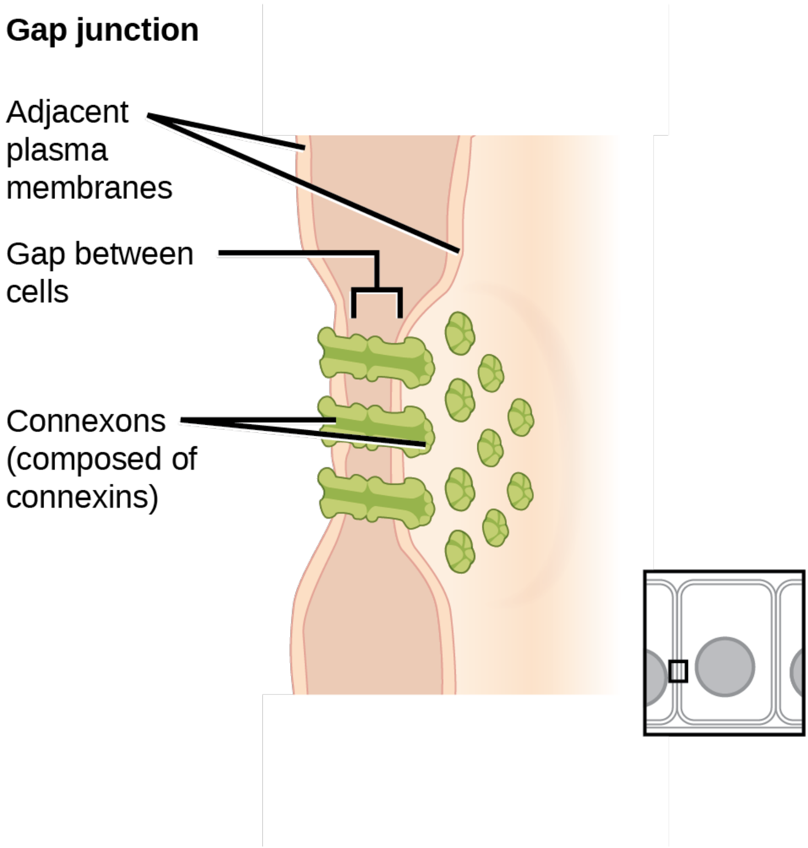
Also only in animal cells are **desmosomes**, which act like spot welds between adjacent epithelial cells ([link]). Cadherins, short proteins in the plasma membrane connect to intermediate filaments to

create desmosomes. The cadherins connect two adjacent cells and maintain the cells in a sheet-like formation in organs and tissues that stretch, like the skin, heart, and muscles.



Gap Junctions

Gap junctions in animal cells are like plasmodesmata in plant cells in that they are channels between adjacent cells that allow for transporting ions, nutrients, and other substances that enable cells to communicate ([link]). Structurally, however, gap junctions and plasmodesmata differ.



Gap junctions develop when a set of six proteins (connexins) in the plasma membrane arrange themselves in an elongated donut-like configuration - a connexon. When the connexon's pores (“doughnut holes”) in adjacent animal cells align, a channel between the two cells forms. Gap junctions are particularly important in cardiac muscle. The electrical signal for the muscle to contract passes efficiently through gap junctions, allowing the heart muscle cells to contract in tandem.

Link to Learning

To conduct a virtual microscopy lab and review the parts of a cell, work through the steps of this interactive assignment.

Section Summary

Animal cells communicate via their extracellular matrices and are connected to each other via tight junctions, desmosomes, and gap junctions. Plant cells are connected and communicate with each other via plasmodesmata.

When protein receptors on the plasma membrane's surface of an animal cell bind to a substance in the extracellular matrix, a chain of reactions begins that changes activities taking place within the cell. Plasmodesmata are channels between adjacent plant cells, while gap junctions are channels between adjacent animal cells. However, their structures are quite different. A tight junction is a watertight seal between two adjacent cells, while a desmosome acts like a spot weld.

Review Questions

Which of the following are only in plant cells?

1. gap junctions
2. desmosomes
3. plasmodesmata
4. tight junctions

C

The key components of desmosomes are cadherins and _______.

1. actin
2. microfilaments
3. intermediate filaments
4. microtubules

C

Diseased animal cells may produce molecules that activate death cascades to kill the cells in a controlled manner. Why would neighboring healthy cells also die?

1. The death molecule is passed through desmosomes.
2. The death molecule is passed through

plasmodesmata.
3. The death molecule disrupts the extracellular matrix.
4. The death molecule passes through gap junctions.

D

Critical Thinking Questions

How does the structure of a plasmodesma differ from that of a gap junction?

They differ because plant cell walls are rigid. Plasmodesmata, which a plant cell needs for transportation and communication, are able to allow movement of really large molecules. Gap junctions are necessary in animal cells for transportation and communication.

Explain how the extracellular matrix functions.

The extracellular matrix functions in support and attachment for animal tissues. It also functions in the healing and growth of the

tissue.

Pathogenic *E. coli* have recently been shown to degrade tight junction proteins during infection. How would this provide an advantage to the bacteria?

E. coli infections generally cause food poisoning, meaning that the invading bacteria cross from the lumen of the gut into the rest of the body. Tight junctions hold the epithelial layer that lines the digestive tract together so that the material that crosses into the body is tightly regulated. One way *E. coli* can avoid this regulation is to destroy the tight junctions so that it can enter the body between the epithelial cells, rather than having to go through the cells.

Glossary

desmosome
: linkages between adjacent epithelial cells that form when cadherins in the plasma membrane attach to intermediate filaments

extracellular matrix
: material secreted from animal or fungal cells that provides mechanical protection and

anchoring for the cells in the tissue

gap junction
: channel between two adjacent animal cells that allows ions, nutrients, and low molecular weight substances to pass between cells, enabling the cells to communicate

plasmodesma
: (plural = plasmodesmata) channel that passes between adjacent plant cells' cell walls, connects their cytoplasm, and allows transporting of materials from cell to cell

tight junction
: protein adherence that creates a firm seal between two adjacent animal cells

Introduction

class="introduction" Despite its seeming hustle and bustle, Grand Central Station functions with a high level of organization: People and objects move from one location to another, they cross or are contained within certain boundaries, and they provide a constant flow as part of larger activity. Analogously, a plasma membrane’s functions involve movement within the cell and across boundaries' activities. (credit: modification of work by Randy Le’Moine)

The plasma membrane, the cell membrane, has many functions, but the most basic one is to define the cell's borders and keep the cell functional. The plasma membrane is selectively permeable. This means that the membrane allows some materials to freely enter or leave the cell, while other materials cannot move freely, but require a specialized structure, and occasionally, even energy investment for crossing.

Components and Structure

By the end of this section, you will be able to do the following:

- Understand the cell membrane fluid mosaic model
- Describe phospholipid, protein, and carbohydrate functions in membranes
- Discuss membrane fluidity

A cell’s plasma membrane defines the cell, outlines its borders, and determines the nature of its interaction with its environment (see [link] for a summary). Cells exclude some substances, take in others, and excrete still others, all in controlled quantities. The plasma membrane must be very flexible to allow certain cells, such as red and white blood cells, to change shape as they pass through narrow capillaries. These are the more obvious plasma membrane functions. In addition, the plasma membrane's surface carries markers that allow cells to recognize one another, which is vital for tissue and organ formation during early development, and which later plays a role in the immune response's “self” versus “non-self” distinction.

Among the most sophisticated plasma membrane functions is the ability for complex, integral proteins, receptors to transmit signals. These proteins act both as extracellular input receivers and as intracellular processing activators. These

membrane receptors provide extracellular attachment sites for effectors like hormones and growth factors, and they activate intracellular response cascades when their effectors are bound. Occasionally, viruses hijack receptors (HIV, human immunodeficiency virus, is one example) that use them to gain entry into cells, and at times, the genes encoding receptors become mutated, causing the signal transduction process to malfunction with disastrous consequences.

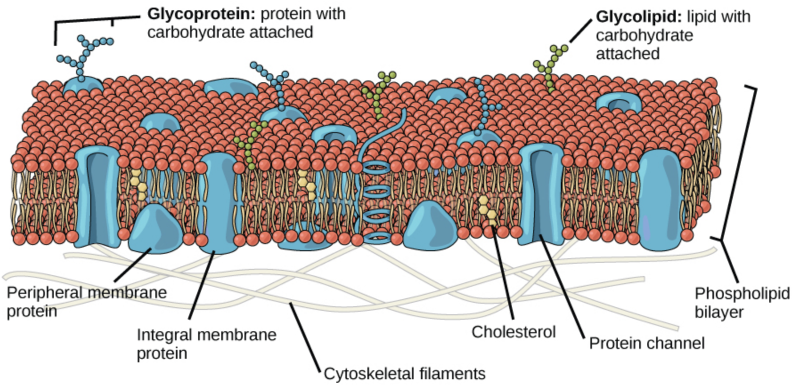
The plasma membrane fluid mosaic model describes the plasma membrane as a fluid combination of phospholipids, cholesterol, and proteins. Carbohydrates attached to lipids (glycolipids) and to proteins (glycoproteins) extend from the membrane's outward-facing surface. A hydrophilic head and two hydrophobic tails comprise this phospholipid molecule. The hydrophilic head group consists of a phosphate-containing group attached to a glycerol molecule. The hydrophobic tails, each containing either a saturated or an unsaturated fatty acid, are long hydrocarbon chains. In an aqueous solution, phospholipids usually arrange themselves with their polar heads facing outward and their hydrophobic tails facing inward. (credit: modification of work by Mariana Ruiz Villareal) Integral membrane proteins may have one or more alpha-helices that span the membrane (examples 1 and 2), or they may have beta-sheets that span the membrane (example 3). (credit: “Foobar”/Wikimedia Commons)

Fluid Mosaic Model

Scientists identified the plasma membrane in the 1890s, and its chemical components in 1915. The principal components they identified were lipids and proteins. In 1935, Hugh Davson and James Danielli proposed the plasma membrane's structure. This was the first model that others in the scientific community widely accepted. It was based on the plasma membrane's “railroad track” appearance in early electron micrographs. Davson and Danielli theorized that the plasma membrane's structure resembles a sandwich. They made the analogy of proteins to bread, and lipids to the filling. In the 1950s, advances in microscopy, notably transmission electron microscopy (TEM), allowed researchers to see that the plasma membrane's core consisted of a double, rather than a single, layer. In 1972, S.J. Singer and Garth L. Nicolson proposed a new model that provides microscopic observations and better explains plasma membrane function.

The explanation, the **fluid mosaic model**, has evolved somewhat over time, but it still best accounts for plasma membrane structure and function as we now understand them. The fluid mosaic model describes the plasma membrane structure as a mosaic of components—including phospholipids, cholesterol, proteins, and carbohydrates—that gives the membrane a fluid character. Plasma membranes range from 5 to 10

nm in thickness. For comparison, human red blood cells, visible via light microscopy, are approximately 8 µm wide, or approximately 1,000 times wider than a plasma membrane. The membrane does look a bit like a sandwich ([link]).



A plasma membrane's principal components are lipids (phospholipids and cholesterol), proteins, and carbohydrates attached to some of the lipids and proteins. A phospholipid is a molecule consisting of glycerol, two fatty acids, and a phosphate-linked head group. Cholesterol, another lipid comprised of four fused carbon rings, is situated alongside the phospholipids in the membrane's core. The protein, lipid, and carbohydrate proportions in the plasma membrane vary with cell type, but for a typical human cell, protein accounts for about 50 percent of the composition by mass, lipids (of all types) account for about 40 percent, and carbohydrates comprise the remaining 10 percent. However, protein and lipid concentration varies with different cell membranes. For example, myelin, an outgrowth of specialized cells' membrane that insulates the

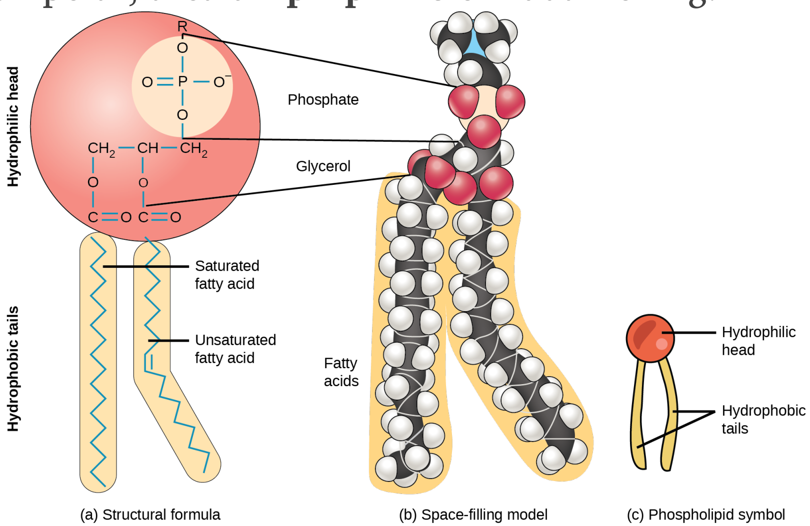
peripheral nerves' axons, contains only 18 percent protein and 76 percent lipid. The mitochondrial inner membrane contains 76 percent protein and only 24 percent lipid. The plasma membrane of human red blood cells is 30 percent lipid. Carbohydrates are present only on the plasma membrane's exterior surface and are attached to proteins, forming **glycoproteins**, or attached to lipids, forming **glycolipids**.

Phospholipids

The membrane's main fabric comprises amphiphilic, phospholipid molecules. The **hydrophilic** or “water-loving” areas of these molecules (which look like a collection of balls in an artist’s rendition of the model) ([link]) are in contact with the aqueous fluid both inside and outside the cell. **Hydrophobic**, or water-hating molecules, tend to be non-polar. They interact with other non-polar molecules in chemical reactions, but generally do not interact with polar molecules. When placed in water, hydrophobic molecules tend to form a ball or cluster. The phospholipids' hydrophilic regions form hydrogen bonds with water and other polar molecules on both the cell's exterior and interior. Thus, the membrane surfaces that face the cell's interior and exterior are hydrophilic. In contrast, the cell membrane's interior is hydrophobic and will not interact with water. Therefore, phospholipids form an excellent two-layer cell membrane that separates fluid within

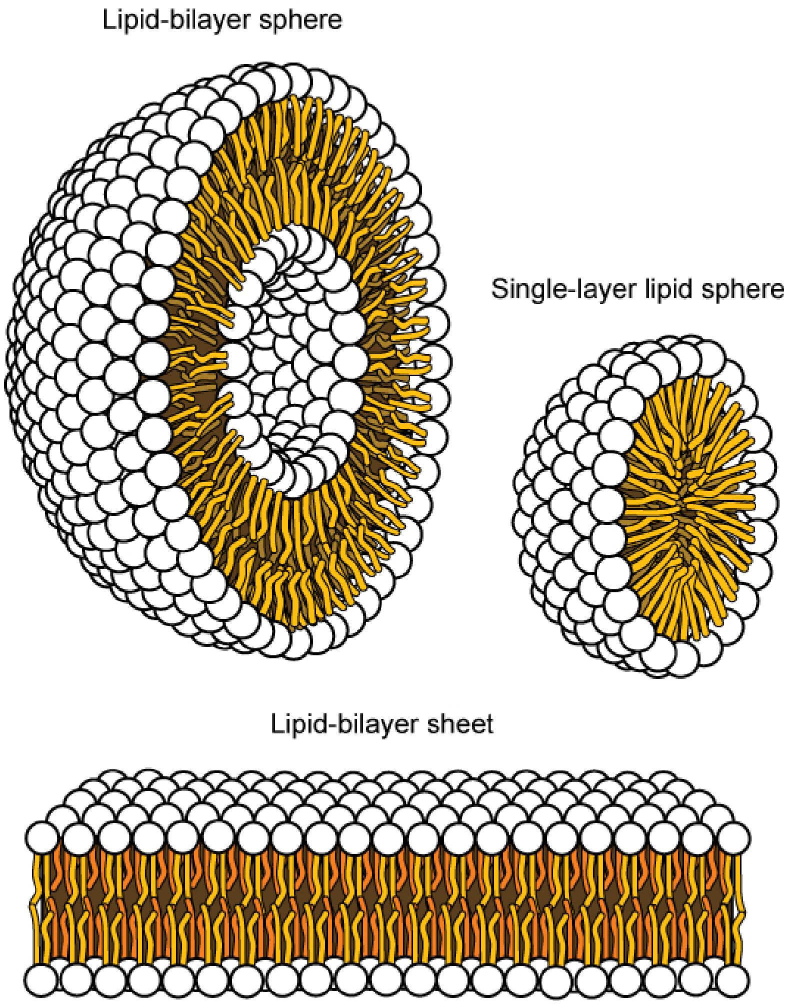
the cell from the fluid outside the cell.

A phospholipid molecule ([link]) consists of a three-carbon glycerol backbone with two fatty acid molecules attached to carbons 1 and 2, and a phosphate-containing group attached to the third carbon. This arrangement gives the overall molecule a head area (the phosphate-containing group), which has a polar character or negative charge, and a tail area (the fatty acids), which has no charge. The head can form hydrogen bonds, but the tail cannot. Scientists call a molecule with a positively or negatively charged area and an uncharged, or non-polar, area **amphiphilic** or “dual-loving.”



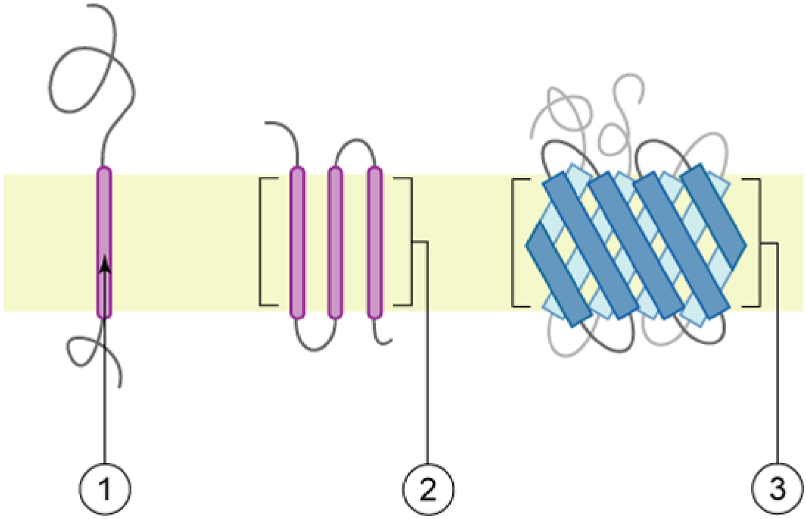
This characteristic is vital to the plasma membrane's structure because, in water, phospholipids arrange themselves with their hydrophobic tails facing each other and their hydrophilic heads facing out. In this way, they form a lipid bilayer—a double layered

phospholipid barrier that separates the water and other materials on one side from the water and other materials on the other side. Phosopholipids heated in an aqueous solution usually spontaneously form small spheres or droplets (micelles or liposomes), with their hydrophilic heads forming the exterior and their hydrophobic tails on the inside ([link]).



Proteins

Proteins comprise the plasma membranes' second major component. **Integral proteins**, or integrins, as their name suggests, integrate completely into the membrane structure, and their hydrophobic membrane-spanning regions interact with the phospholipid bilayer's hydrophobic region ([link]). Single-pass integral membrane proteins usually have a hydrophobic transmembrane segment that consists of 20–25 amino acids. Some span only part of the membrane—associating with a single layer—while others stretch from one side to the other, and are exposed on either side. Up to 12 single protein segments comprise some complex proteins, which are extensively folded and embedded in the membrane ([link]). This protein type has a hydrophilic region or regions, and one or several mildly hydrophobic regions. This arrangement of protein regions orients the protein alongside the phospholipids, with the protein's hydrophobic region adjacent to the phosopholipids' tails and the protein's hydrophilic region or regions protruding from the membrane and in contact with the cytosol or extracellular fluid.



Peripheral proteins are on the membranes' exterior and interior surfaces, attached either to integral proteins or to phospholipids. Peripheral proteins, along with integral proteins, may serve as enzymes, as structural attachments for the cytoskeleton's fibers, or as part of the cell's recognition sites. Scientists sometimes refer to these as "cell-specific" proteins. The body recognizes its own proteins and attacks foreign proteins associated with invasive pathogens.

Carbohydrates

Carbohydrates are the third major plasma membrane component. They are always on the cells' exterior surface and are bound either to proteins (forming glycoproteins) or to lipids (forming glycolipids) ([link]). These carbohydrate chains may consist of 2–60 monosaccharide units and can be

either straight or branched. Along with peripheral proteins, carbohydrates form specialized sites on the cell surface that allow cells to recognize each other. These sites have unique patterns that allow for cell recognition, much the way that the facial features unique to each person allow individuals to recognize him or her. This recognition function is very important to cells, as it allows the immune system to differentiate between body cells (“self”) and foreign cells or tissues (“non-self”). Similar glycoprotein and glycolipid types are on the surfaces of viruses and may change frequently, preventing immune cells from recognizing and attacking them.

We collectively refer to these carbohydrates on the cell's exterior surface—the carbohydrate components of both glycoproteins and glycolipids—as the glycocalyx (meaning “sugar coating”). The glycocalyx is highly hydrophilic and attracts large amounts of water to the cell's surface. This aids in the cell's interaction with its watery environment and in the cell’s ability to obtain substances dissolved in the water. As we discussed above, the glycocalyx is also important for cell identification, self/non-self determination, and embryonic development, and is used in cell to cell attachments to form tissues.

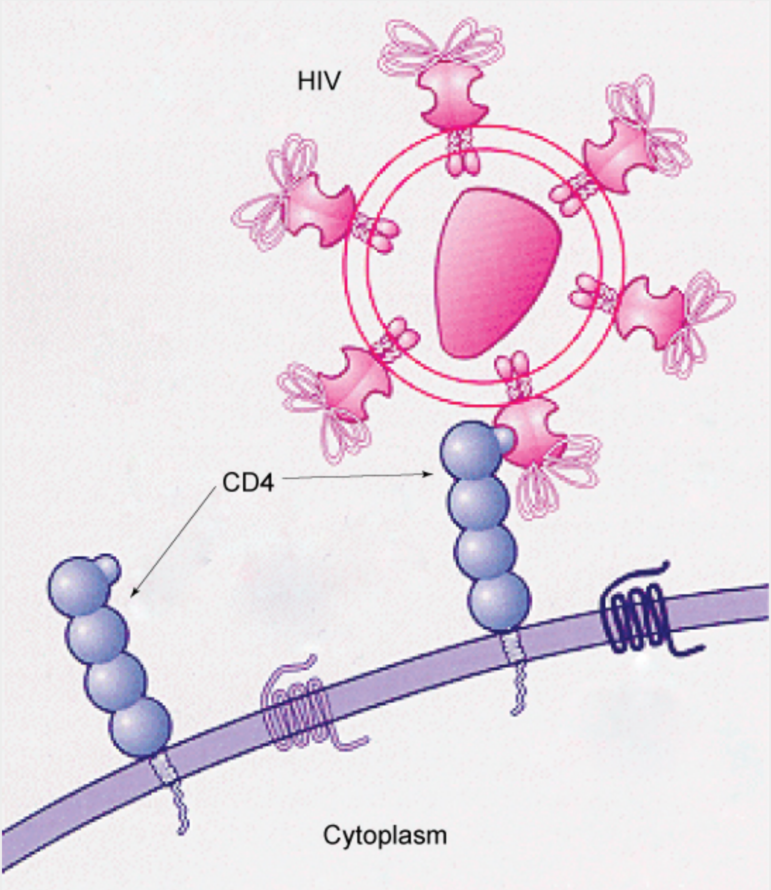
Evolution Connection

How Viruses Infect Specific Organs

Glycoprotein and glycolipid patterns on the cells' surfaces give many viruses an opportunity for infection. HIV and hepatitis viruses infect only specific organs or cells in the human body. HIV is able to penetrate the plasma membranes of a subtype of lymphocytes called T-helper cells, as well as some monocytes and central nervous system cells. The hepatitis virus attacks liver cells. These viruses are able to invade these cells, because the cells have binding sites on their surfaces that are specific to and compatible with certain viruses ([link]). Other recognition sites on the virus’s surface interact with the human immune system, prompting the body to produce antibodies. Antibodies are made in response to the antigens or proteins associated with invasive pathogens, or in response to foreign cells, such as might occur with an organ transplant. These same sites serve as places for antibodies to attach and either destroy or inhibit the virus' activity. Unfortunately, these recognition sites on HIV change at a rapid rate because of mutations, making an effective vaccine against the virus very difficult, as the virus evolves and adapts. A person infected with HIV will quickly develop different populations, or variants, of the virus that differences in these recognition sites distinguish. This rapid change of surface markers decreases the effectiveness of the person’s immune system in attacking the virus, because the antibodies will not recognize the surface patterns'

new variations. In the case of HIV, the problem is compounded because the virus specifically infects and destroys cells involved in the immune response, further incapacitating the host.

HIV binds to the CD4 receptor, a glycoprotein on T cell surfaces. (credit: modification of work by NIH, NIAID)



Membrane Fluidity

The membrane's mosaic characteristic helps to illustrate its nature. The integral proteins and lipids exist in the membrane as separate but loosely attached molecules. These resemble the separate, multicolored tiles of a mosaic picture, and they float, moving somewhat with respect to one another. The membrane is not like a balloon, however, that can expand and contract; rather, it is fairly rigid and can burst if penetrated or if a cell takes in too much water. However, because of its mosaic nature, a very fine needle can easily penetrate a plasma membrane without causing it to burst, and the membrane will flow and self-seal when one extracts the needle.

The membrane's mosaic characteristics explain some but not all of its fluidity. There are two other factors that help maintain this fluid characteristic. One factor is the nature of the phospholipids themselves. In their saturated form, the fatty acids in phospholipid tails are saturated with bound hydrogen atoms. There are no double bonds between adjacent carbon atoms. This results in tails that are relatively straight. In contrast, unsaturated fatty acids do not contain a maximal number of hydrogen atoms, but they do contain some double bonds between adjacent carbon atoms. A double bond results in a bend in the carbon string of approximately 30 degrees ([link]).

Thus, if decreasing temperatures compress saturated fatty acids with their straight tails, they press in on each other, making a dense and fairly rigid membrane. If unsaturated fatty acids are compressed, the “kinks” in their tails elbow adjacent phospholipid molecules away, maintaining some space between the phospholipid molecules. This “elbow room” helps to maintain fluidity in the membrane at temperatures at which membranes with saturated fatty acid tails in their phospholipids would “freeze” or solidify. The membrane's relative fluidity is particularly important in a cold environment. A cold environment usually compresses membranes comprised largely of saturated fatty acids, making them less fluid and more susceptible to rupturing. Many organisms (fish are one example) are capable of adapting to cold environments by changing the proportion of unsaturated fatty acids in their membranes in response to lower temperature.

Link to Learning
Visit this site to see animations of the membranes' fluidity and mosaic quality.

Animals have an additional membrane constituent that assists in maintaining fluidity. Cholesterol,

which lies alongside the phospholipids in the membrane, tends to dampen temperature effects on the membrane. Thus, this lipid functions as a buffer, preventing lower temperatures from inhibiting fluidity and preventing increased temperatures from increasing fluidity too much. Thus, cholesterol extends, in both directions, the temperature range in which the membrane is appropriately fluid and consequently functional. Cholesterol also serves other functions, such as organizing clusters of transmembrane proteins into lipid rafts.

Plasma Membrane Components and Functions

Component	**Location**
Phospholipid	Main membrane fabric
Cholesterol	Attached between phospholipids and between the two phospholipid layers
Integral proteins (for example, integrins)	Embedded within the phospholipid layer(s); may or may not penetrate through both layers
Peripheral proteins	On the phospholipid bilayer's inner or outer

	surface; not embedded within the phospholipids
Carbohydrates (components of glycoproteins and glycolipids)	Generally attached to proteins on the outside membrane layer

Career Connection

Immunologist

The variations in peripheral proteins and carbohydrates that affect a cell’s recognition sites are of prime interest in immunology. In developing vaccines, researchers have been able to conquer many infectious diseases, such as smallpox, polio, diphtheria, and tetanus.

Immunologists are the physicians and scientists who research and develop vaccines, as well as treat and study allergies or other immune problems. Some immunologists study and treat autoimmune problems (diseases in which a person’s immune system attacks his or her own cells or tissues, such as lupus) and immunodeficiencies, whether acquired (such as acquired immunodeficiency syndrome, or AIDS) or hereditary (such as severe combined immunodeficiency, or SCID). Immunologists also help treat organ transplantation patients, who must have their immune systems suppressed so that their bodies will not reject a transplanted organ. Some

immunologists work to understand natural immunity and the effects of a person’s environment on it. Others work on questions about how the immune system affects diseases such as cancer. In the past, researchers did not understand the importance of having a healthy immune system in preventing cancer.

To work as an immunologist, one must have a PhD or MD. In addition, immunologists undertake at least two to three years of training in an accredited program and must pass the American Board of Allergy and Immunology exam. Immunologists must possess knowledge of the human body's function as they relate to issues beyond immunization, and knowledge of pharmacology and medical technology, such as medications, therapies, test materials, and surgical procedures.

Section Summary

Modern scientists refer to the plasma membrane as the fluid mosaic model. A phospholipid bilayer comprises the plasma membrane, with hydrophobic, fatty acid tails in contact with each other. The membrane's landscape is studded with proteins, some which span the membrane. Some of these proteins serve to transport materials into or out of

the cell. Carbohydrates are attached to some of the proteins and lipids on the membrane's outward-facing surface, forming complexes that function to identify the cell to other cells. The membrane's fluid nature is due to temperature, fatty acid tail configuration (some kinked by double bonds), cholesterol presence embedded in the membrane, and the mosaic nature of the proteins and protein-carbohydrate combinations, which are not firmly fixed in place. Plasma membranes enclose and define the cells' borders. Not static, they are dynamic and constantly in flux.

Review Questions

Which plasma membrane component can be either found on its surface or embedded in the membrane structure?

1. protein
2. cholesterol
3. carbohydrate
4. phospholipid

A

Which characteristic of a phospholipid

contributes to the fluidity of the membrane?

1. its head
2. cholesterol
3. a saturated fatty acid tail
4. double bonds in the fatty acid tail

D

What is the primary function of carbohydrates attached to the exterior of cell membranes?

1. identification of the cell
2. flexibility of the membrane
3. strengthening the membrane
4. channels through membrane

A

A scientist compares the plasma membrane composition of an animal from the Mediterranean coast with one from the Mojave Desert. Which hypothesis is most likely to be correct?

1. The cells from the Mediterranean coast animal will have more fluid plasma membranes.

2. The cells from the Mojave Desert animal will have a higher cholesterol concentration in the plasma membranes.
3. The cells' plasma membranes will be indistinguishable.
4. The cells from the Mediterranean coast animal will have a higher glycoprotein content, while the cells from the Mojave Desert animal will have a higher lipoprotein content.

B

Critical Thinking Questions

Why is it advantageous for the cell membrane to be fluid in nature?

The fluid characteristic of the cell membrane allows greater flexibility to the cell than it would if the membrane were rigid. It also allows the motion of membrane components, required for some types of membrane transport.

Why do phospholipids tend to spontaneously

orient themselves into something resembling a membrane?

The hydrophobic, nonpolar regions must align with each other in order for the structure to have minimal potential energy and, consequently, higher stability. The fatty acid tails of the phospholipids cannot mix with water, but the phosphate “head” of the molecule can. Thus, the head orients to water, and the tail to other lipids.

How can a cell use an extracellular peripheral protein as the receptor to transmit a signal into the cell?

Peripheral proteins can bind to other molecules in the extracellular space. However, they cannot directly transmit a signal to the inside of the cell since they do not have a transmembrane domain (region that goes through the plasma membrane to the inside of the cell). They must associate with integral membrane proteins in order to pass the signal to the inside of the cell.

Glossary

amphiphilic
: molecule possessing a polar or charged area and a nonpolar or uncharged area capable of interacting with both hydrophilic and hydrophobic environments

fluid mosaic model
: describes the plasma membrane's structure as a mosaic of components including phospholipids, cholesterol, proteins, glycoproteins, and glycolipids (sugar chains attached to proteins or lipids, respectively), resulting in a fluid character (fluidity)

glycolipid
: combination of carbohydrates and lipids

glycoprotein
: combination of carbohydrates and proteins

hydrophilic
: molecule with the ability to bond with water; “water-loving”

hydrophobic
: molecule that does not have the ability to bond with water; “water-hating”

integral protein
: protein integrated into the membrane structure that interacts extensively with the membrane lipids' hydrocarbon chains and

often spans the membrane

peripheral protein
protein at the plasma membrane's surface either on its exterior or interior side

Passive Transport

By the end of this section, you will be able to do the following:

- Explain why and how passive transport occurs
- Understand the osmosis and diffusion processes
- Define tonicity and its relevance to passive transport

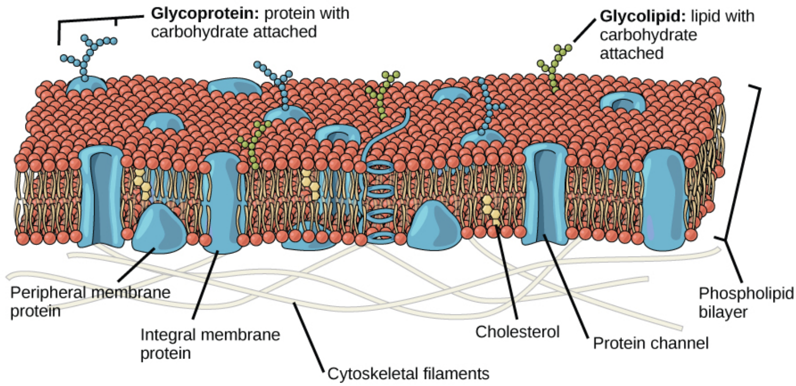
Plasma membranes must allow certain substances to enter and leave a cell, and prevent some harmful materials from entering and some essential materials from leaving. In other words, plasma membranes are **selectively permeable—**they allow some substances to pass through, but not others. If they were to lose this selectivity, the cell would no longer be able to sustain itself, and it would be destroyed. Some cells require larger amounts of specific substances. They must have a way of obtaining these materials from extracellular fluids. This may happen passively, as certain materials move back and forth, or the cell may have special mechanisms that facilitate transport. Some materials are so important to a cell that it spends some of its energy, hydrolyzing adenosine triphosphate (ATP), to obtain these materials. Red blood cells use some of their energy doing just that. Most cells spend the majority of their energy to maintain an imbalance of sodium and potassium ions between the cell's interior and exterior, as well as on protein synthesis.

The most direct forms of membrane transport are passive. **Passive transport** is a naturally occurring phenomenon and does not require the cell to exert any of its energy to accomplish the movement. In passive transport, substances move from an area of higher concentration to an area of lower concentration. A physical space in which there is a single substance concentration range has a **concentration gradient**.

The plasma membrane's exterior surface is not identical to its interior surface.

Selective Permeability

Plasma membranes are asymmetric: the membrane's interior is not identical to its exterior. There is a considerable difference between the array of phospholipids and proteins between the two leaflets that form a membrane. On the membrane's interior, some proteins serve to anchor the membrane to cytoskeleton's fibers. There are peripheral proteins on the membrane's exterior that bind extracellular matrix elements. Carbohydrates, attached to lipids or proteins, are also on the plasma membrane's exterior surface. These carbohydrate complexes help the cell bind required substances in the extracellular fluid. This adds considerably to plasma membrane's selective nature ([link]).



Recall that plasma membranes are amphiphilic: They have hydrophilic and hydrophobic regions. This characteristic helps move some materials through the membrane and hinders the movement of others. Non-polar and lipid-soluble material with a low molecular weight can easily slip through the membrane's hydrophobic lipid core. Substances such as the fat-soluble vitamins A, D, E, and K readily pass through the plasma membranes in the digestive tract and other tissues. Fat-soluble drugs and hormones also gain easy entry into cells and readily transport themselves into the body’s tissues and organs. Oxygen and carbon dioxide molecules have no charge and pass through membranes by simple diffusion.

Polar substances present problems for the membrane. While some polar molecules connect easily with the cell's outside, they cannot readily pass through the plasma membrane's lipid core. Additionally, while small ions could easily slip through the spaces in the membrane's mosaic, their

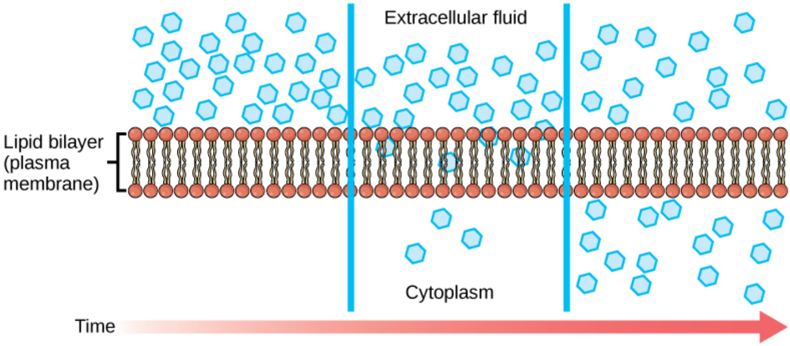
charge prevents them from doing so. Ions such as sodium, potassium, calcium, and chloride must have special means of penetrating plasma membranes. Simple sugars and amino acids also need the help of various transmembrane proteins (channels) to transport themselves across plasma membranes.

Diffusion through a permeable membrane moves a substance from a high concentration area (extracellular fluid, in this case) down its concentration gradient (into the cytoplasm). (credit: modification of work by Mariana Ruiz Villareal)

Diffusion

Diffusion is a passive process of transport. A single substance moves from a high concentration to a low concentration area until the concentration is equal across a space. You are familiar with diffusion of substances through the air. For example, think about someone opening a bottle of ammonia in a room filled with people. The ammonia gas is at its highest concentration in the bottle. Its lowest concentration is at the room's edges. The ammonia vapor will diffuse, or spread away, from the bottle, and gradually, increasingly more people will smell the ammonia as it spreads. Materials move within the cell’s cytosol by diffusion, and certain materials move through the plasma membrane by diffusion ([link]). Diffusion expends no energy. On the contrary, concentration gradients are a form of potential energy, which dissipates as the gradient is

eliminated.



Each separate substance in a medium, such as the extracellular fluid, has its own concentration gradient, independent of other materials' concentration gradients. In addition, each substance will diffuse according to that gradient. Within a system, there will be different diffusion rates of various substances in the medium.

Factors That Affect Diffusion

Molecules move constantly in a random manner, at a rate that depends on their mass, their environment, and the amount of thermal energy they possess, which in turn is a function of temperature. This movement accounts for molecule diffusion through whatever medium in which they are localized. A substance moves into any space available to it until it evenly distributes itself throughout. After a substance has diffused completely through a space, removing its concentration gradient, molecules will still move

around in the space, but there will be no *net* movement of the number of molecules from one area to another. We call this lack of a concentration gradient in which the substance has no net movement dynamic equilibrium. While diffusion will go forward in the presence of a substance's concentration gradient, several factors affect the diffusion rate.

- Extent of the concentration gradient: The greater the difference in concentration, the more rapid the diffusion. The closer the distribution of the material gets to equilibrium, the slower the diffusion rate.
- Mass of the molecules diffusing: Heavier molecules move more slowly; therefore, they diffuse more slowly. The reverse is true for lighter molecules.
- Temperature: Higher temperatures increase the energy and therefore the molecules' movement, increasing the diffusion rate. Lower temperatures decrease the molecules' energy, thus decreasing the diffusion rate.
- Solvent density: As the density of a solvent increases, the diffusion rate decreases. The molecules slow down because they have a more difficult time passing through the denser medium. If the medium is less dense, diffusion increases. Because cells primarily use diffusion to move materials within the cytoplasm, any increase in the cytoplasm’s density will inhibit

the movement of the materials. An example of this is a person experiencing dehydration. As the body’s cells lose water, the diffusion rate decreases in the cytoplasm, and the cells’ functions deteriorate. Neurons tend to be very sensitive to this effect. Dehydration frequently leads to unconsciousness and possibly coma because of the decrease in diffusion rate within the cells.

- Solubility: As we discussed earlier, nonpolar or lipid-soluble materials pass through plasma membranes more easily than polar materials, allowing a faster diffusion rate.
- Surface area and plasma membrane thickness: Increased surface area increases the diffusion rate; whereas, a thicker membrane reduces it.
- Distance travelled: The greater the distance that a substance must travel, the slower the diffusion rate. This places an upper limitation on cell size. A large, spherical cell will die because nutrients or waste cannot reach or leave the cell's center, respectively. Therefore, cells must either be small in size, as in the case of many prokaryotes, or be flattened, as with many single-celled eukaryotes.

A variation of diffusion is the process of filtration. In filtration, material moves according to its concentration gradient through a membrane. Sometimes pressure enhances the diffusion rate, causing the substances to filter more rapidly. This

occurs in the kidney, where blood pressure forces large amounts of water and accompanying dissolved substances, or **solutes**, out of the blood and into the renal tubules. The diffusion rate in this instance is almost totally dependent on pressure. One of the effects of high blood pressure is the appearance of protein in the urine, which abnormally high pressure "squeezes through".

Facilitated transport moves substances down their concentration gradients. They may cross the plasma membrane with the aid of channel proteins. (credit: modification of work by Mariana Ruiz Villareal)

Some substances are able to move down their concentration gradient across the plasma membrane with the aid of carrier proteins. Carrier proteins change shape as they move molecules across the membrane. (credit: modification of work by Mariana Ruiz Villareal)

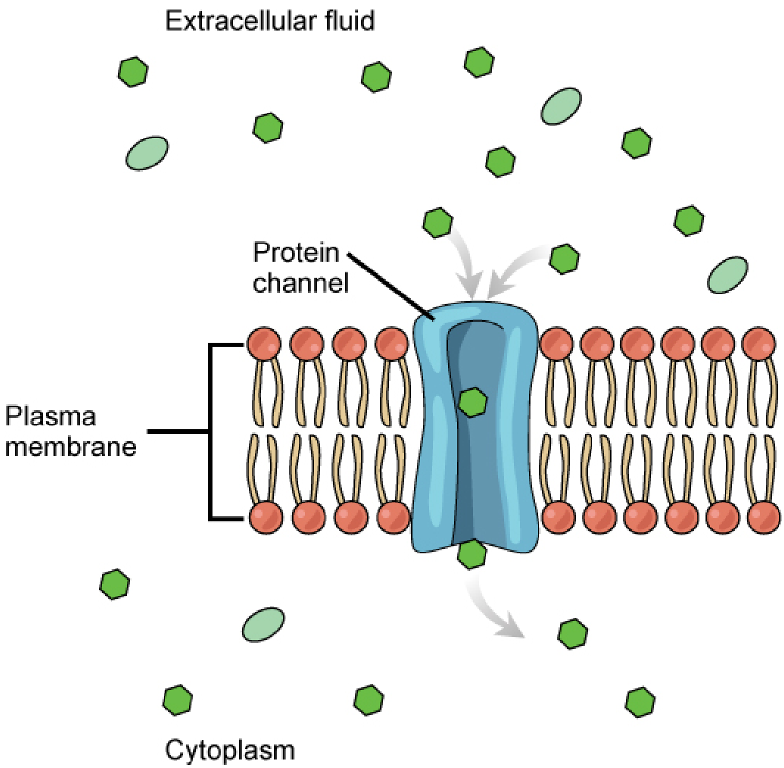
Facilitated transport

In **facilitated transport**, or facilitated diffusion, materials diffuse across the plasma membrane with the help of membrane proteins. A concentration gradient exists that would allow these materials to diffuse into the cell without expending cellular energy. However, these materials are polar molecule ions that the cell membrane's hydrophobic parts repel. Facilitated transport proteins shield these materials from the membrane's repulsive force, allowing them to diffuse into the cell.

The transported material first attaches to protein or glycoprotein receptors on the plasma membrane's exterior surface. This allows removal of material from the extracellular fluid that the cell needs. The substances then pass to specific integral proteins that facilitate their passage. Some of these integral proteins are collections of beta-pleated sheets that form a pore or channel through the phospholipid bilayer. Others are carrier proteins which bind with the substance and aid its diffusion through the membrane.

Channels

The integral proteins involved in facilitated transport are **transport proteins**, and they function as either channels for the material or carriers. In both cases, they are transmembrane proteins. Channels are specific for the transported substance. **Channel proteins** have hydrophilic domains exposed to the intracellular and extracellular fluids. In addition, they have a hydrophilic channel through their core that provides a hydrated opening through the membrane layers ([link]). Passage through the channel allows polar compounds to avoid the plasma membrane's nonpolar central layer that would otherwise slow or prevent their entry into the cell. **Aquaporins** are channel proteins that allow water to pass through the membrane at a very high rate.

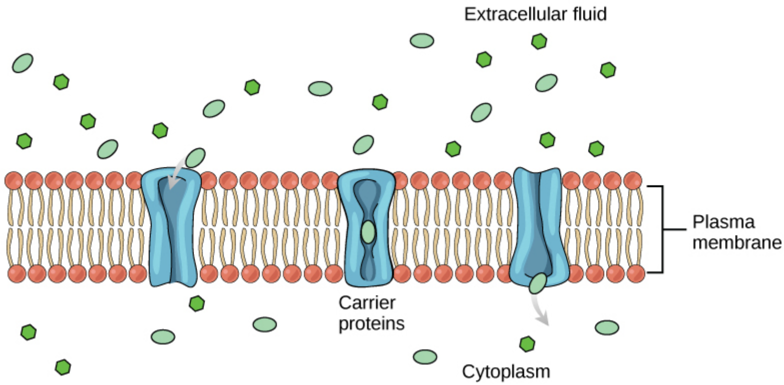


Channel proteins are either open at all times or they are “gated,” which controls the channel's opening. When a particular ion attaches to the channel protein it may control the opening, or other mechanisms or substances may be involved. In some tissues, sodium and chloride ions pass freely through open channels; whereas, in other tissues a gate must open to allow passage. An example of this occurs in the kidney, where there are both channel forms in different parts of the renal tubules. Cells involved in transmitting electrical impulses, such as nerve and muscle cells, have gated channels for sodium, potassium, and calcium in their

membranes. Opening and closing these channels changes the relative concentrations on opposing sides of the membrane of these ions, resulting in facilitating electrical transmission along membranes (in the case of nerve cells) or in muscle contraction (in the case of muscle cells).

Carrier Proteins

Another type of protein embedded in the plasma membrane is a **carrier protein**. This aptly named protein binds a substance and, thus triggers a change of its own shape, moving the bound molecule from the cell's outside to its interior ([link]). Depending on the gradient, the material may move in the opposite direction. Carrier proteins are typically specific for a single substance. This selectivity adds to the plasma membrane's overall selectivity. Scientists poorly understand the exact mechanism for the change of shape. Proteins can change shape when their hydrogen bonds are affected, but this may not fully explain this mechanism. Each carrier protein is specific to one substance, and there are a finite number of these proteins in any membrane. This can cause problems in transporting enough material for the cell to function properly. When all of the proteins are bound to their ligands, they are saturated and the rate of transport is at its maximum. Increasing the concentration gradient at this point will not result in an increased transport rate.



An example of this process occurs in the kidney. In one part, the kidney filters glucose, water, salts, ions, and amino acids that the body requires. This filtrate, which includes glucose, then reabsorbs in another part of the kidney. Because there are only a finite number of carrier proteins for glucose, if more glucose is present than the proteins can handle, the excess is not transported and the body excretes this through urine. In a diabetic individual, the term is “spilling glucose into the urine.” A different group of carrier proteins, glucose transport proteins, or GLUTs, are involved in transporting glucose and other hexose sugars through plasma membranes within the body.

Channel and carrier proteins transport material at different rates. Channel proteins transport much more quickly than carrier proteins. Channel proteins facilitate diffusion at a rate of tens of millions of molecules per second; whereas, carrier proteins work at a rate of a thousand to a million molecules per second.

In osmosis, water always moves from an area of higher water concentration to one of lower concentration. In the diagram, the solute cannot pass through the selectively permeable membrane, but the water can.

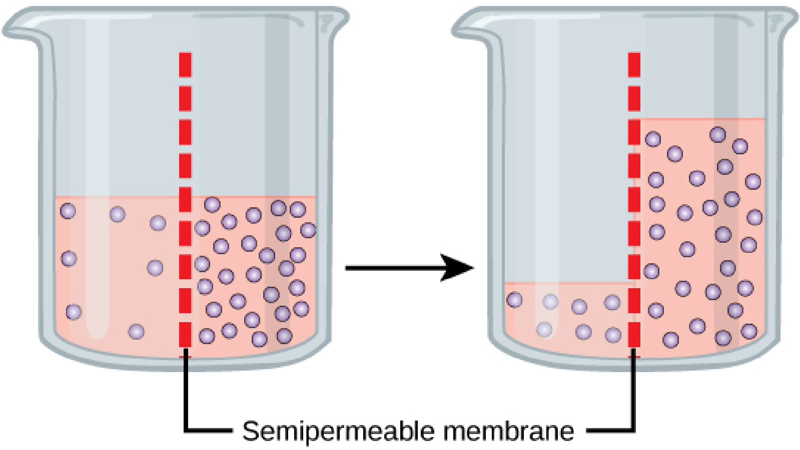
Osmosis

Osmosis is the movement of water through a semipermeable membrane according to the water's concentration gradient across the membrane, which is inversely proportional to the solutes' concentration. While diffusion transports material across membranes and within cells, osmosis transports *only water* across a membrane and the membrane limits the solutes' diffusion in the water. Not surprisingly, the aquaporins that facilitate water movement play a large role in osmosis, most prominently in red blood cells and the membranes of kidney tubules.

Mechanism

Osmosis is a special case of diffusion. Water, like other substances, moves from an area of high concentration to one of low concentration. An obvious question is what makes water move at all? Imagine a beaker with a semipermeable membrane separating the two sides or halves ([link]). On both sides of the membrane the water level is the same,

but there are different dissolved substance concentrations, or **solute**, that cannot cross the membrane (otherwise the solute crossing the membrane would balance concentrations on each side). If the solution's volume on both sides of the membrane is the same, but the solute's concentrations are different, then there are different amounts of water, the solvent, on either side of the membrane.



To illustrate this, imagine two full water glasses. One has a single teaspoon of sugar in it; whereas, the second one contains one-quarter cup of sugar. If the total volume of the solutions in both cups is the same, which cup contains more water? Because the large sugar amount in the second cup takes up much more space than the teaspoon of sugar in the first cup, the first cup has more water in it.

Returning to the beaker example, recall that it has a solute mixture on either side of the membrane. A

principle of diffusion is that the molecules move around and will spread evenly throughout the medium if they can. However, only the material capable of getting through the membrane will diffuse through it. In this example, the solute cannot diffuse through the membrane, but the water can. Water has a concentration gradient in this system. Thus, water will diffuse down its concentration gradient, crossing the membrane to the side where it is less concentrated. This diffusion of water through the membrane—osmosis—will continue until the water's concentration gradient goes to zero or until the water's hydrostatic pressure balances the osmotic pressure. Osmosis proceeds constantly in living systems.

Tonicity

Tonicity describes how an extracellular solution can change a cell's volume by affecting osmosis. A solution's tonicity often directly correlates with the solution's osmolarity. **Osmolarity** describes the solution's total solute concentration. A solution with low osmolarity has a greater number of water molecules relative to the number of solute particles. A solution with high osmolarity has fewer water molecules with respect to solute particles. In a situation in which a membrane permeable to water, though not to the solute separates two different osmolarities, water will move from the membrane's

side with lower osmolarity (and more water) to the side with higher osmolarity (and less water). This effect makes sense if you remember that the solute cannot move across the membrane, and thus the only component in the system that can move—the water—moves along its own concentration gradient. An important distinction that concerns living systems is that osmolarity measures the number of particles (which may be molecules) in a solution. Therefore, a solution that is cloudy with cells may have a lower osmolarity than a solution that is clear, if the second solution contains more dissolved molecules than there are cells.

Hypotonic Solutions

Scientists use three terms—hypotonic, isotonic, and hypertonic—to relate the cell's osmolarity to the extracellular fluid's osmolarity that contains the cells. In a **hypotonic** situation, the extracellular fluid has lower osmolarity than the fluid inside the cell, and water enters the cell. (In living systems, the point of reference is always the cytoplasm, so the prefix *hypo-* means that the extracellular fluid has a lower solute concentration, or a lower osmolarity, than the cell cytoplasm.) It also means that the extracellular fluid has a higher water concentration in the solution than does the cell. In this situation, water will follow its concentration gradient and enter the cell.

Hypertonic Solutions

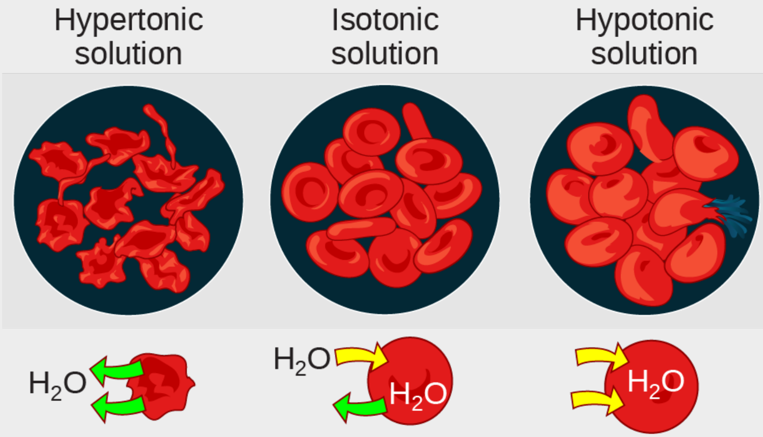
As for a **hypertonic** solution, the prefix *hyper-* refers to the extracellular fluid having a higher osmolarity than the cell's cytoplasm; therefore, the fluid contains less water than the cell does. Because the cell has a relatively higher water concentration, water will leave the cell.

Isotonic Solutions

In an **isotonic** solution, the extracellular fluid has the same osmolarity as the cell. If the cell's osmolarity matches that of the extracellular fluid, there will be no net movement of water into or out of the cell, although water will still move in and out. Blood cells and plant cells in hypertonic, isotonic, and hypotonic solutions take on characteristic appearances ([link]).

Visual Connection

Osmotic pressure changes red blood cells' shape in hypertonic, isotonic, and hypotonic solutions. (credit: Mariana Ruiz Villareal)



A doctor injects a patient with what the doctor thinks is an isotonic saline solution. The patient dies, and an autopsy reveals that many red blood cells have been destroyed. Do you think the solution the doctor injected was really isotonic?

Link to Learning

For a video illustrating the diffusion process in solutions, visit this site.

The turgor pressure within a plant cell depends on the solution's tonicity in which it is bathed. (credit: modification of work by Mariana Ruiz Villareal) Without adequate water, the plant on the left has lost turgor pressure, visible in its wilting. Watering the plant (right) will restore the turgor pressure. (credit: Victor M. Vicente Selvas) A paramecium’s

contractile vacuole, here visualized using bright field light microscopy at 480x magnification, continuously pumps water out of the organism's body to keep it from bursting in a hypotonic medium. (credit: modification of work by NIH; scale-bar data from Matt Russell)

Tonicity in Living Systems

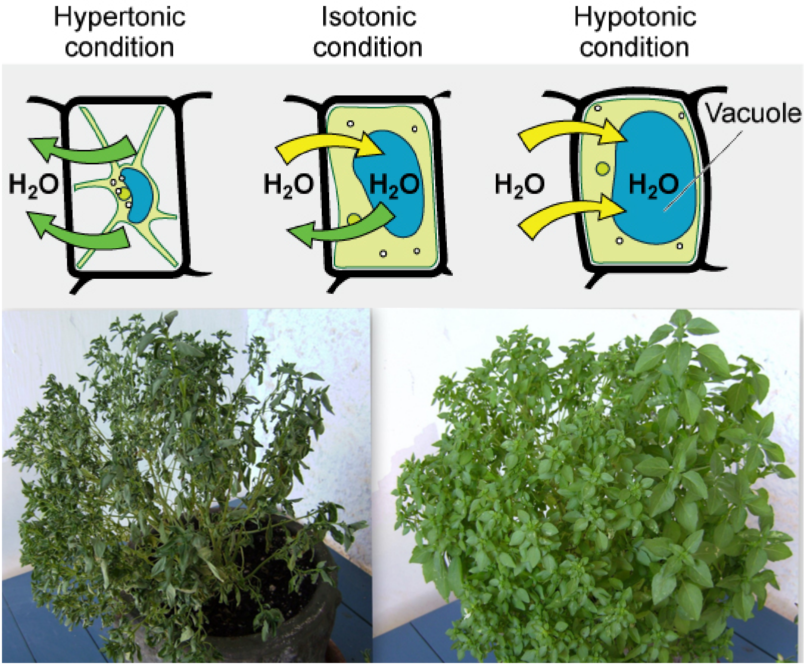
In a hypotonic environment, water enters a cell, and the cell swells. In an isotonic condition, the relative solute and solvent concentrations are equal on both membrane sides. There is no net water movement; therefore, there is no change in the cell's size. In a hypertonic solution, water leaves a cell and the cell shrinks. If either the hypo- or hyper- condition goes to excess, the cell's functions become compromised, and the cell may be destroyed.

A red blood cell will burst, or lyse, when it swells beyond the plasma membrane's capability to expand. Remember, the membrane resembles a mosaic, with discrete spaces between the molecules comprising it. If the cell swells, and the spaces between the lipids and proteins become too large, the cell will break apart.

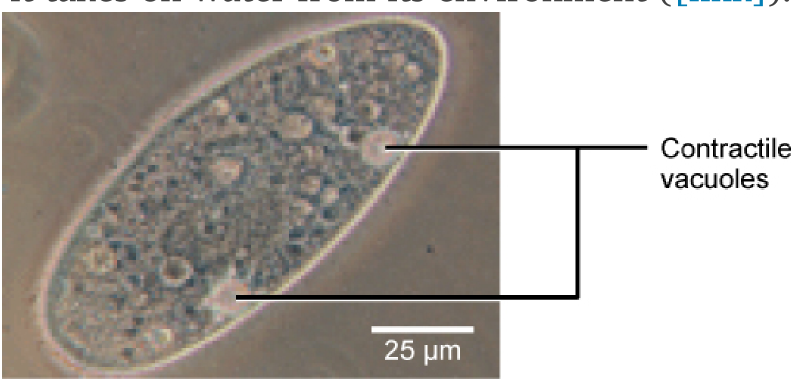
In contrast, when excessive water amounts leave a red blood cell, the cell shrinks, or crenates. This has the effect of concentrating the solutes left in the cell, making the cytosol denser and interfering with

diffusion within the cell. The cell’s ability to function will be compromised and may also result in the cell's death.

Various living things have ways of controlling the effects of osmosis—a mechanism we call osmoregulation. Some organisms, such as plants, fungi, bacteria, and some protists, have cell walls that surround the plasma membrane and prevent cell lysis in a hypotonic solution. The plasma membrane can only expand to the cell wall's limit, so the cell will not lyse. The cytoplasm in plants is always slightly hypertonic to the cellular environment, and water will always enter a cell if water is available. This water inflow produces turgor pressure, which stiffens the plant's cell walls ([link]). In nonwoody plants, turgor pressure supports the plant. Conversly, if you do not water the plant, the extracellular fluid will become hypertonic, causing water to leave the cell. In this condition, the cell does not shrink because the cell wall is not flexible. However, the cell membrane detaches from the wall and constricts the cytoplasm. We call this **plasmolysis**. Plants lose turgor pressure in this condition and wilt ([link]).



Tonicity is a concern for all living things. For example, paramecia and amoebas, which are protists that lack cell walls, have contractile vacuoles. This vesicle collects excess water from the cell and pumps it out, keeping the cell from lysing as it takes on water from its environment ([link]).



Many marine invertebrates have internal salt levels matched to their environments, making them isotonic with the water in which they live. Fish, however, must spend approximately five percent of their metabolic energy maintaining osmotic homeostasis. Freshwater fish live in an environment that is hypotonic to their cells. These fish actively take in salt through their gills and excrete diluted urine to rid themselves of excess water. Saltwater fish live in the reverse environment, which is hypertonic to their cells, and they secrete salt through their gills and excrete highly concentrated urine.

In vertebrates, the kidneys regulate the water amount in the body. Osmoreceptors are specialized cells in the brain that monitor solute concentration in the blood. If the solute levels increase beyond a certain range, a hormone releases that slows water loss through the kidney and dilutes the blood to safer levels. Animals also have high albumin concentrations, which the liver produces, in their blood. This protein is too large to pass easily through plasma membranes and is a major factor in controlling the osmotic pressures applied to tissues.

Section Summary

The passive transport forms, diffusion and osmosis, move materials of small molecular weight across

membranes. Substances diffuse from high to lower concentration areas, and this process continues until the substance evenly distributes itself in a system. In solutions containing more than one substance, each molecule type diffuses according to its own concentration gradient, independent of other substances diffusing. Many factors can affect the diffusion rate, such as concentration gradient, diffusing, particle sizes, and the system's temperature.

In living systems, the plasma membrane mediates substances diffusing in and out of cells. Some materials diffuse readily through the membrane, but others are hindered and only can pass through due to specialized proteins such as channels and transporters. The chemistry of living things occurs in aqueous solutions, and balancing the concentrations of those solutions is an ongoing problem. In living systems, diffusing some substances would be slow or difficult without membrane proteins that facilitate transport.

Visual Connection Questions

[link] A doctor injects a patient with what the doctor thinks is an isotonic saline solution. The patient dies, and an autopsy reveals that many

red blood cells have been destroyed. Do you think the solution the doctor injected was really isotonic?

[link] No, it must have been hypotonic as a hypotonic solution would cause water to enter the cells, thereby making them burst.

Review Questions

Water moves via osmosis _______.

1. throughout the cytoplasm
2. from an area with a high concentration of other solutes to a lower one
3. from an area with a high concentration of water to one of lower concentration
4. from an area with a low concentration of water to higher concentration

---

C

The principal force driving movement in diffusion is the ________.

1. temperature
2. particle size
3. concentration gradient
4. membrane surface area

C

What problem is faced by organisms that live in fresh water?

1. Their bodies tend to take in too much water.
2. They have no way of controlling their tonicity.
3. Only salt water poses problems for animals that live in it.
4. Their bodies tend to lose too much water to their environment.

A

In which situation would passive transport **not** use a transport protein for entry into a cell?

1. water flowing into a hypertonic environment
2. glucose being absorbed from the blood
3. an ion flowing into a nerve cell to create

an electrical potential
4. oxygen moving into a cell after oxygen deprivation

D

Critical Thinking Questions

Discuss why the following affect the rate of diffusion: molecular size, temperature, solution density, and the distance that must be traveled.

Heavy molecules move more slowly than lighter ones. It takes more energy in the medium to move them along. Increasing or decreasing temperature increases or decreases the energy in the medium, affecting molecular movement. The denser a solution is, the harder it is for molecules to move through it, causing diffusion to slow down due to friction. Living cells require a steady supply of nutrients and a steady rate of waste removal. If the distance these substances need to travel is too great, diffusion cannot move nutrients and waste materials efficiently to sustain life.

Why does water move through a membrane?

Water moves through a membrane in osmosis because there is a concentration gradient across the membrane of solute and solvent. The solute cannot effectively move to balance the concentration on both sides of the membrane, so water moves to achieve this balance.

Both of the regular intravenous solutions administered in medicine, normal saline and lactated Ringer's solution, are isotonic. Why is this important?

Injection of isotonic solutions ensures that there will be no perturbation of the osmotic balance, and no water taken from tissues or added to them from the blood.

Describe two ways that decreasing temperature would affect the rate of diffusion of molecules across a cell's plasma membrane.

Decreasing temperature will decrease the kinetic energy in the system. A lower temperature means less energy in the molecules, so they will move at a slower speed.

Lowering temperature also decreases the kinetic energy of the molecules in the plasma membrane, compressing them together. This increases the density of the plasma membrane, which slows diffusion into the cell.

A cell develops a mutation in its potassium channels that prevents the ions from leaving the cell. If the cell's aquaporins are still active, what will happen to the cell? Be sure to describe the tonicity and osmolarity of the cell.

Without functional potassium channels, the potassium ions that are pumped into the cell will accumulate. This increases the osmolarity inside the cell, creating a hypotonic solution. Since the plasma membrane is still selectively permeable to water by the aquaporins, water will flow into the cell. If the potassium concentration is high enough, enough water will eventually flow into the cell to lyse it.

Glossary

aquaporin
: channel protein that allows water through the membrane at a very high rate

carrier protein

membrane protein that moves a substance across the plasma membrane by changing its own shape

channel protein
: membrane protein that allows a substance to pass through its hollow core across the plasma membrane

concentration gradient
: area of high concentration adjacent to an area of low concentration

diffusion
: passive transport process of low-molecular weight material according to its concentration gradient

facilitated transport
: process by which material moves down a concentration gradient (from high to low concentration) using integral membrane proteins

hypertonic
: situation in which extracellular fluid has a higher osmolarity than the fluid inside the cell, resulting in water moving out of the cell

hypotonic
: situation in which extracellular fluid has a lower osmolarity than the fluid inside the

cell, resulting in water moving into the cell

isotonic
: situation in which the extracellular fluid has the same osmolarity as the fluid inside the cell, resulting in no net water movement into or out of the cell

osmolarity
: total amount of substances dissolved in a specific amount of solution

osmosis
: transport of water through a semipermeable membrane according to the water's concentration gradient across the membrane that results from the presence of solute that cannot pass through the membrane

passive transport
: method of transporting material through a membrane that does not require energy

plasmolysis
: detaching the cell membrane from the cell wall and constricting the cell membrane when a plant cell is in a hypertonic solution

selectively permeable
: membrane characteristic that allows some substances through

solute
: substance dissolved in a liquid to form a solution

tonicity
: amount of solute in a solution

transport protein
: membrane protein that facilitates a substance's passage across a membrane by binding it

Active Transport
By the end of this section, you will be able to do the following:

- Understand how electrochemical gradients affect ions
- Distinguish between primary active transport and secondary active transport

Active transport mechanisms require the cell’s energy, usually in the form of adenosine triphosphate (ATP). If a substance must move into the cell against its concentration gradient—that is, if the substance's concentration inside the cell is greater than its concentration in the extracellular fluid (and vice versa)—the cell must use energy to move the substance. Some active transport mechanisms move small-molecular weight materials, such as ions, through the membrane. Other mechanisms transport much larger molecules. A uniporter carries one molecule or ion. A symporter carries two different molecules or ions, both in the same direction. An antiporter also carries two different molecules or ions, but in different directions. (credit: modification of work by “Lupask”/Wikimedia Commons)

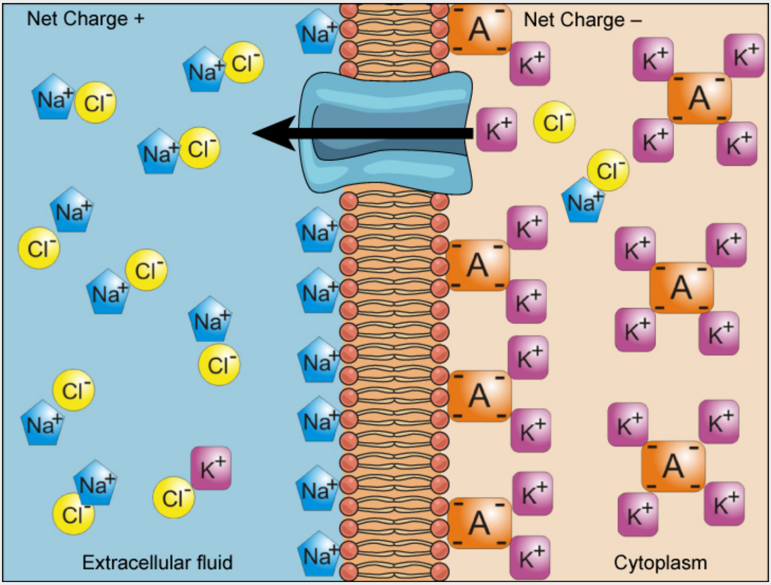
Electrochemical Gradient

We have discussed simple concentration gradients—

a substance's differential concentrations across a space or a membrane—but in living systems, gradients are more complex. Because ions move into and out of cells and because cells contain proteins that do not move across the membrane and are mostly negatively charged, there is also an electrical gradient, a difference of charge, across the plasma membrane. The interior of living cells is electrically negative with respect to the extracellular fluid in which they are bathed, and at the same time, cells have higher concentrations of potassium (K^+) and lower concentrations of sodium (Na^+) than the extracellular fluid. Thus in a living cell, the concentration gradient of Na^+ tends to drive it into the cell, and its electrical gradient (a positive ion) also drives it inward to the negatively charged interior. However, the situation is more complex for other elements such as potassium. The electrical gradient of K^+, a positive ion, also drives it into the cell, but the concentration gradient of K^+ drives K^+ *out* of the cell ([link]). We call the combined concentration gradient and electrical charge that affects an ion its **electrochemical gradient**.

Visual Connection

Electrochemical gradients arise from the combined effects of concentration gradients and electrical gradients. Structures labeled A represent proteins. (credit: “Synaptitude”/Wikimedia Commons)



Injecting a potassium solution into a person’s blood is lethal. This is how capital punishment and euthanasia subjects die. Why do you think a potassium solution injection is lethal?

Moving Against a Gradient

To move substances against a concentration or electrochemical gradient, the cell must use energy. This energy comes from ATP generated through the cell’s metabolism. Active transport mechanisms, or **pumps**, work against electrochemical gradients. Small substances constantly pass through plasma membranes. Active transport maintains concentrations of ions and other substances that living cells require in the face of these passive

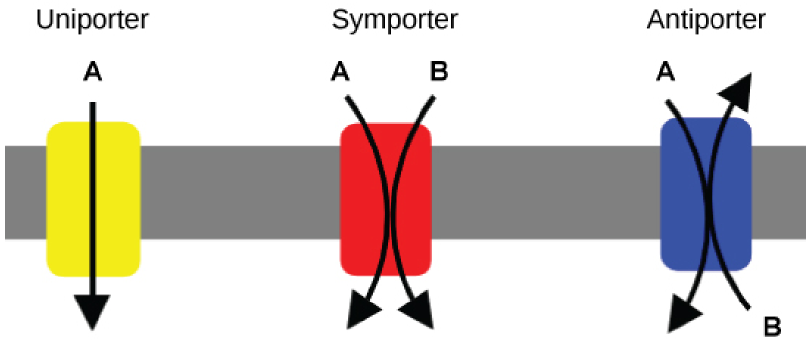
movements. A cell may spend much of its metabolic energy supply maintaining these processes. (A red blood cell uses most of its metabolic energy to maintain the imbalance between exterior and interior sodium and potassium levels that the cell requires.) Because active transport mechanisms depend on a cell’s metabolism for energy, they are sensitive to many metabolic poisons that interfere with the ATP supply.

Two mechanisms exist for transporting small-molecular weight material and small molecules. **Primary active transport** moves ions across a membrane and creates a difference in charge across that membrane, which is directly dependent on ATP. **Secondary active transport** does not directly require ATP: instead, it is the movement of material due to the electrochemical gradient established by primary active transport.

Carrier Proteins for Active Transport

An important membrane adaption for active transport is the presence of specific carrier proteins or pumps to facilitate movement: there are three protein types or **transporters** ([link]). A **uniporter** carries one specific ion or molecule. A **symporter** carries two different ions or molecules, both in the same direction. An **antiporter** also carries two different ions or molecules, but in different directions. All of these transporters can also

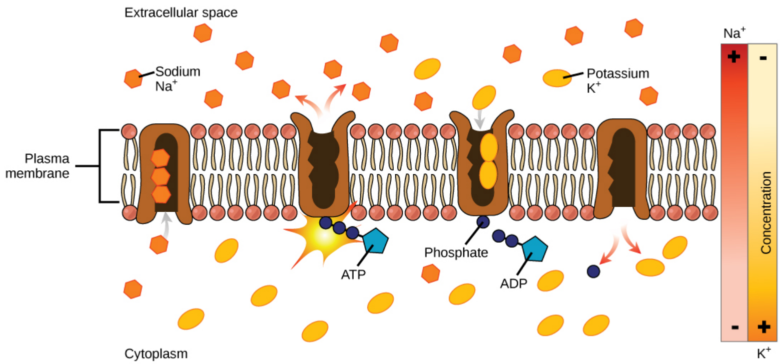
transport small, uncharged organic molecules like glucose. These three types of carrier proteins are also in facilitated diffusion, but they do not require ATP to work in that process. Some examples of pumps for active transport are Na^+-K^+ ATPase, which carries sodium and potassium ions, and H^+-K^+ ATPase, which carries hydrogen and potassium ions. Both of these are antiporter carrier proteins. Two other carrier proteins are Ca^{2+} ATPase and H^+ ATPase, which carry only calcium and only hydrogen ions, respectively. Both are pumps.



Primary active transport moves ions across a membrane, creating an electrochemical gradient (electrogenic transport). (credit: modification of work by Mariana Ruiz Villareal)

Primary Active Transport

The primary active transport that functions with the active transport of sodium and potassium allows secondary active transport to occur. The second transport method is still active because it depends on using energy as does primary transport ([link]).



One of the most important pumps in animal cells is the sodium-potassium pump (Na^+-K^+ ATPase), which maintains the electrochemical gradient (and the correct concentrations of Na^+ and K^+) in living cells. The sodium-potassium pump moves K^+ into the cell while moving Na^+ out at the same time, at a ratio of three Na^+ for every two K^+ ions moved in. The Na^+-K^+ ATPase exists in two forms, depending on its orientation to the cell's interior or exterior and its affinity for either sodium or potassium ions. The process consists of the following six steps.

1. With the enzyme oriented towards the cell's interior, the carrier has a high affinity for sodium ions. Three ions bind to the protein.
2. The protein carrier hydrolyzes ATP and a low-energy phosphate group attaches to it.
3. As a result, the carrier changes shape and reorients itself towards the membrane's exterior. The protein's affinity for sodium decreases and the three sodium ions leave the

carrier.

4. The shape change increases the carrier's affinity for potassium ions, and two such ions attach to the protein. Subsequently, the low-energy phosphate group detaches from the carrier.
5. With the phosphate group removed and potassium ions attached, the carrier protein repositions itself towards the cell's interior.
6. The carrier protein, in its new configuration, has a decreased affinity for potassium, and the two ions moves into the cytoplasm. The protein now has a higher affinity for sodium ions, and the process starts again.

Several things have happened as a result of this process. At this point, there are more sodium ions outside the cell than inside and more potassium ions inside than out. For every three sodium ions that move out, two potassium ions move in. This results in the interior being slightly more negative relative to the exterior. This difference in charge is important in creating the conditions necessary for the secondary process. The sodium-potassium pump is, therefore, an **electrogenic pump** (a pump that creates a charge imbalance), creating an electrical imbalance across the membrane and contributing to the membrane potential.

Link to Learning

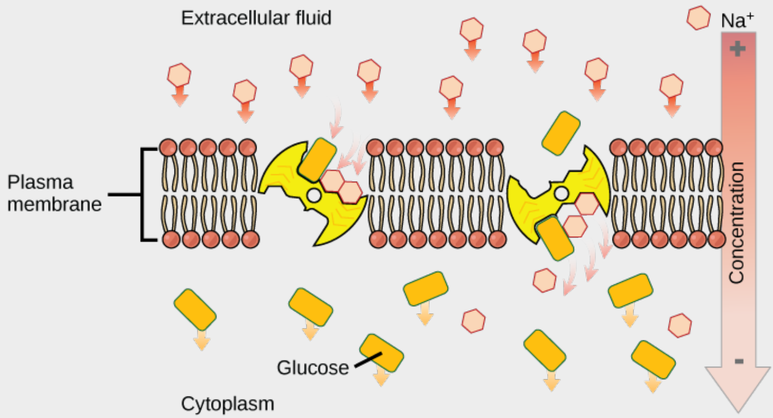
Watch this video to see an active transport simulation in a sodium-potassium ATPase.

Secondary Active Transport (Co-transport)

Secondary active transport brings sodium ions, and possibly other compounds, into the cell. As sodium ion concentrations build outside of the plasma membrane because of the primary active transport process, this creates an electrochemical gradient. If a channel protein exists and is open, the sodium ions will pull through the membrane. This movement transports other substances that can attach themselves to the transport protein through the membrane ([link]). Many amino acids, as well as glucose, enter a cell this way. This secondary process also stores high-energy hydrogen ions in the mitochondria of plant and animal cells in order to produce ATP. The potential energy that accumulates in the stored hydrogen ions translates into kinetic energy as the ions surge through the channel protein ATP synthase, and that energy then converts ADP into ATP.

Visual Connection

An electrochemical gradient, which primary active transport creates, can move other substances against their concentration gradients, a process scientists call co-transport or secondary active transport. (credit: modification of work by Mariana Ruiz Villareal)



If the pH outside the cell decreases, would you expect the amount of amino acids transported into the cell to increase or decrease?

Section Summary

The combined gradient that affects an ion includes its concentration gradient and its electrical gradient. A positive ion, for example, might diffuse into a new area, down its concentration gradient, but if it is

diffusing into an area of net positive charge, its electrical gradient hampers its diffusion. When dealing with ions in aqueous solutions, one must consider electrochemical and concentration gradient combinations, rather than just the concentration gradient alone. Living cells need certain substances that exist inside the cell in concentrations greater than they exist in the extracellular space. Moving substances up their electrochemical gradients requires energy from the cell. Active transport uses energy stored in ATP to fuel this transport. Active transport of small molecular-sized materials uses integral proteins in the cell membrane to move the materials. These proteins are analogous to pumps. Some pumps, which carry out primary active transport, couple directly with ATP to drive their action. In co-transport (or secondary active transport), energy from primary transport can move another substance into the cell and up its concentration gradient.

Visual Connection Questions

[link] Injecting a potassium solution into a person's blood is lethal. Capital punishment and euthanasia utilize this method in their subjects. Why do you think a potassium solution injection is lethal?

[link] Cells typically have a high concentration of potassium in the cytoplasm and are bathed in a high concentration of sodium. Injection of potassium dissipates this electrochemical gradient. In heart muscle, the sodium/potassium potential is responsible for transmitting the signal that causes the muscle to contract. When this potential is dissipated, the signal can't be transmitted, and the heart stops beating. Potassium injections are also used to stop the heart from beating during surgery.

[link] If the pH outside the cell decreases, would you expect the amount of amino acids transported into the cell to increase or decrease?

[link] A decrease in pH means an increase in positively charged $H+$ ions, and an increase in the electrical gradient across the membrane. The transport of amino acids into the cell will increase.

Review Questions

Active transport must function continuously because _______.

1. plasma membranes wear out
2. not all membranes are amphiphilic
3. facilitated transport opposes active transport
4. diffusion is constantly moving solutes in opposite directions

---

D

How does the sodium-potassium pump make the interior of the cell negatively charged?

1. by expelling anions
2. by pulling in anions
3. by expelling more cations than are taken in
4. by taking in and expelling an equal number of cations

C

What is the combination of an electrical gradient and a concentration gradient called?

1. potential gradient
2. electrical potential

3. concentration potential
4. electrochemical gradient

D

Critical Thinking Questions

Where does the cell get energy for active transport processes?

The cell harvests energy from ATP produced by its own metabolism to power active transport processes, such as the activity of pumps.

How does the sodium-potassium pump contribute to the net negative charge of the interior of the cell?

The sodium-potassium pump forces out three (positive) Na^+ ions for every two (positive) K^+ ions it pumps in, thus the cell loses a positive charge at every cycle of the pump.

Glucose from digested food enters intestinal

epithelial cells by active transport. Why would intestinal cells use active transport when most body cells use facilitated diffusion?

Intestinal epithelial cells use active transport to fulfill their specific role as the cells that transfer glucose from the digested food to the bloodstream. Intestinal cells are exposed to an environment with fluctuating glucose levels. Immediately after eating, glucose in the gut lumen will be high, and could accumulate in intestinal cells by diffusion. However, when the gut lumen is empty, glucose levels are higher in the intestinal cells. If glucose moved by facilitated diffusion, this would cause glucose to flow back out of the intestinal cells and into the gut. Active transport proteins ensure that glucose moves into the intestinal cells, and cannot move back into the gut. It also ensures that glucose transport continues to occur even if high levels of glucose are already present in the intestinal cells. This maximizes the amount of energy the body can harvest from food.

The sodium/calcium exchanger (NCX) transports sodium into and calcium out of cardiac muscle cells. Describe why this transporter is classified as secondary active transport.

The NCX moves sodium down its electrochemical gradient into the cell. Since sodium's electrochemical gradient is created by the $Na+/K+$ pump, a transport pump that requires ATP hydrolysis to establish the gradient, the NCX is a secondary active transport process.

Glossary

active transport
: method of transporting material that requires energy

antiporter
: transporter that carries two ions or small molecules in different directions

electrochemical gradient
: a combined electrical and chemical force that produces a gradient

electrogenic pump
: pump that creates a charge imbalance

primary active transport
: active transport that moves ions or small molecules across a membrane and may create a difference in charge across that membrane

pump

active transport mechanism that works against electrochemical gradients

secondary active transport
: movement of material that results from primary active transport to the electrochemical gradient

symporter
: transporter that carries two different ions or small molecules, both in the same direction

transporter
: specific carrier proteins or pumps that facilitate movement

uniporter
: transporter that carries one specific ion or molecule

Bulk Transport
By the end of this section, you will be able to do the following:

- Describe endocytosis, including phagocytosis, pinocytosis, and receptor-mediated endocytosis
- Understand the process of exocytosis

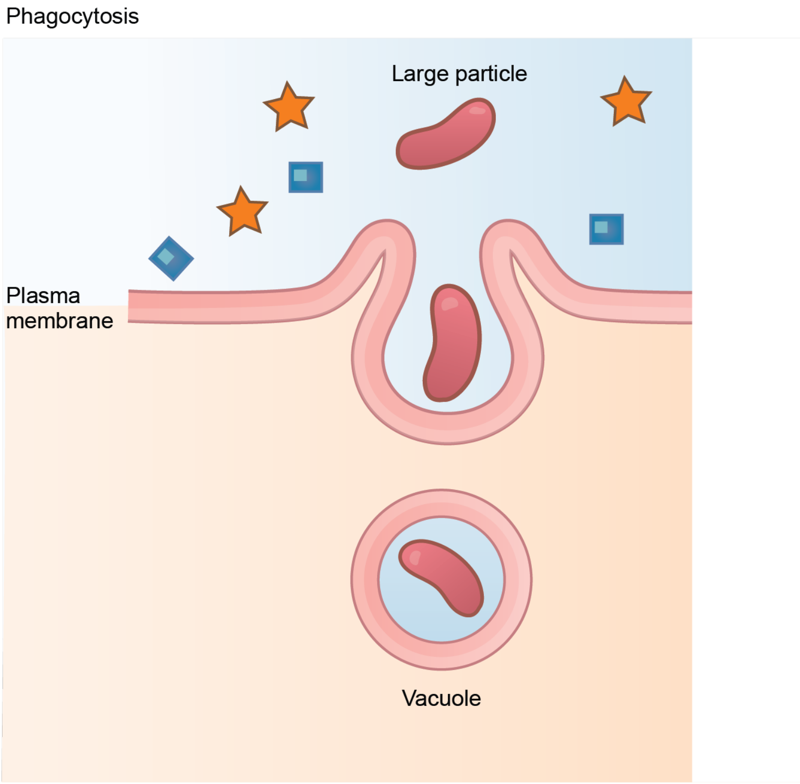
In addition to moving small ions and molecules through the membrane, cells also need to remove and take in larger molecules and particles (see [link] for examples). Some cells are even capable of engulfing entire unicellular microorganisms. You might have correctly hypothesized that when a cell uptakes and releases large particles, it requires energy. A large particle, however, cannot pass through the membrane, even with energy that the cell supplies.
In phagocytosis, the cell membrane surrounds the particle and engulfs it. (credit: modification of work by Mariana Ruiz Villareal) In pinocytosis, the cell membrane invaginates, surrounds a small volume of fluid, and pinches off. (credit: modification of work by Mariana Ruiz Villareal) In receptor-mediated endocytosis, the cell's uptake of substances targets a single type of substance that binds to the receptor on the cell membrane's external surface. (credit: modification of work by Mariana Ruiz Villareal)

Endocytosis

Endocytosis is a type of active transport that moves particles, such as large molecules, parts of cells, and even whole cells, into a cell. There are different endocytosis variations, but all share a common characteristic: the cell's plasma membrane invaginates, forming a pocket around the target particle. The pocket pinches off, resulting in the particle containing itself in a newly created intracellular vesicle formed from the plasma membrane.

Phagocytosis

Phagocytosis (the condition of “cell eating”) is the process by which a cell takes in large particles, such as other cells or relatively large particles. For example, when microorganisms invade the human body, a type of white blood cell, a neutrophil, will remove the invaders through this process, surrounding and engulfing the microorganism, which the neutrophil then destroys ([link]).



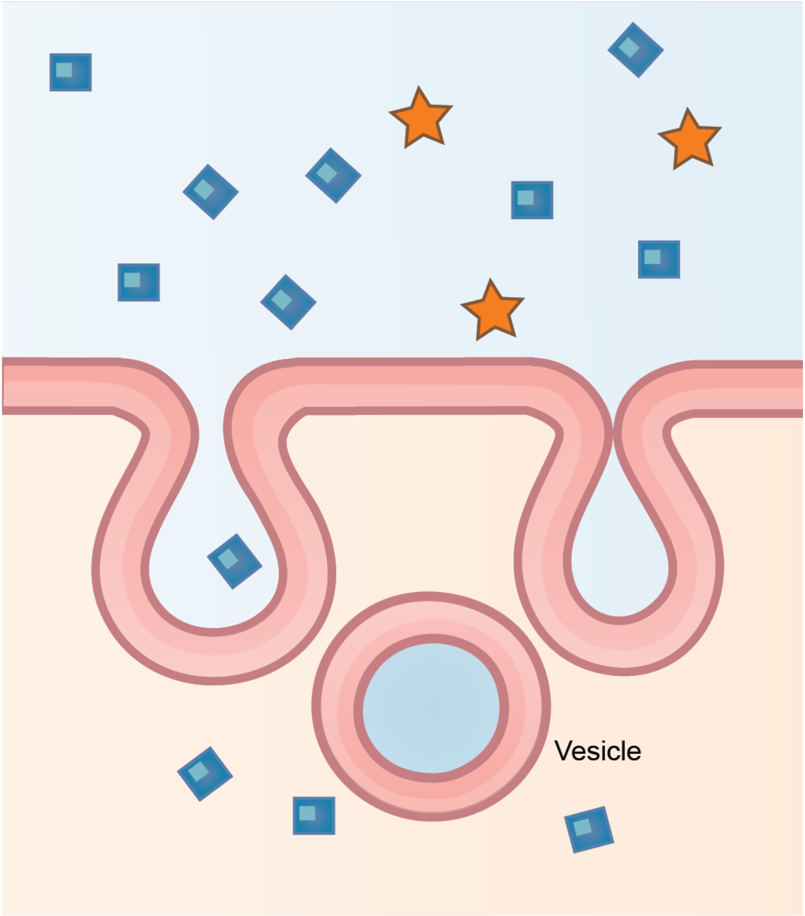
In preparation for phagocytosis, a portion of the plasma membrane's inward-facing surface becomes coated with the protein **clathrin**, which stabilizes this membrane's section. The membrane's coated portion then extends from the cell's body and surrounds the particle, eventually enclosing it. Once the vesicle containing the particle is enclosed within the cell, the clathrin disengages from the membrane and the vesicle merges with a lysosome for breaking down the material in the newly formed compartment (endosome). When accessible nutrients from the vesicular contents' degradation have been extracted, the newly formed endosome

merges with the plasma membrane and releases its contents into the extracellular fluid. The endosomal membrane again becomes part of the plasma membrane.

Pinocytosis

A variation of endocytosis is **pinocytosis**. This literally means “cell drinking”. Discovered by Warren Lewis in 1929, this American embryologist and cell biologist described a process whereby he assumed that the cell was purposefully taking in extracellular fluid. In reality, this is a process that takes in molecules, including water, which the cell needs from the extracellular fluid. Pinocytosis results in a much smaller vesicle than does phagocytosis, and the vesicle does not need to merge with a lysosome ([link]).

Pinocytosis



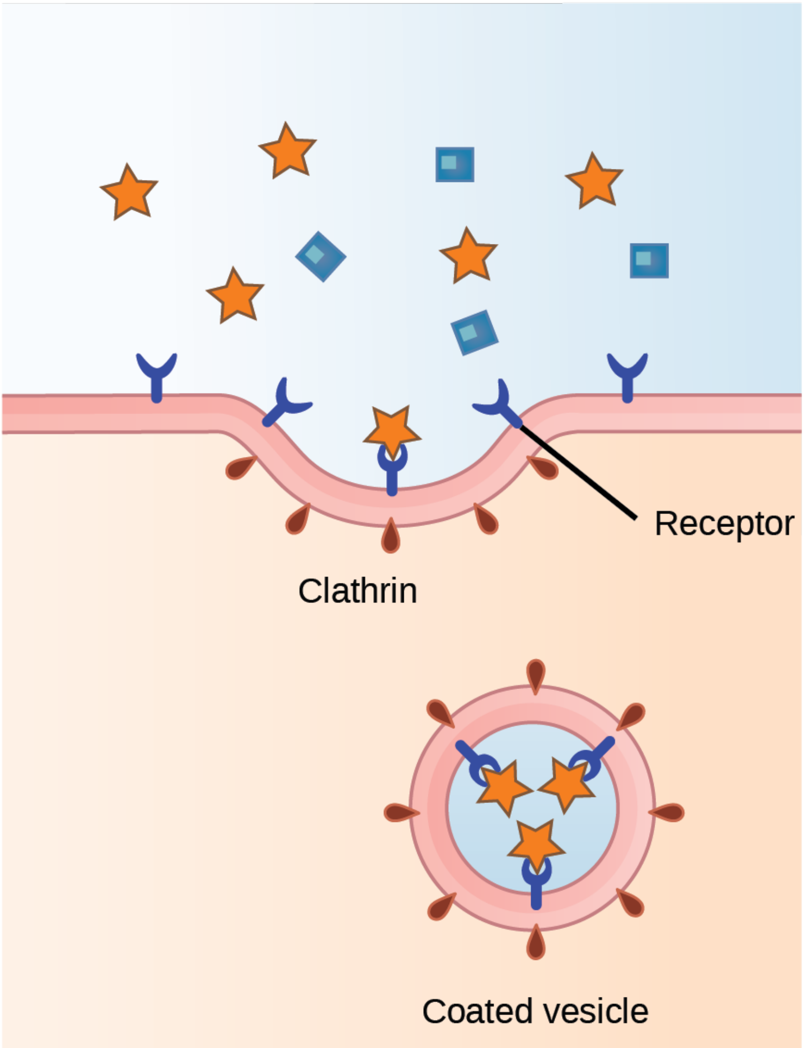
A variation of pinocytosis is **potocytosis**. This process uses a coating protein, **caveolin**, on the plasma membrane's cytoplasmic side, which performs a similar function to clathrin. The cavities in the plasma membrane that form the vacuoles have membrane receptors and lipid rafts in addition to caveolin. The vacuoles or vesicles formed in caveolae (singular caveola) are smaller than those in pinocytosis. Potocytosis brings small molecules into

the cell and transports them through the cell for their release on the other side, a process we call transcytosis.

Receptor-mediated Endocytosis

A targeted variation of endocytosis employs receptor proteins in the plasma membrane that have a specific binding affinity for certain substances ([link]).

Receptor-mediated endocytosis



In **receptor-mediated endocytosis**, as in phagocytosis, clathrin attaches to the plasma membrane's cytoplasmic side. If a compound's uptake is dependent on receptor-mediated endocytosis and the process is ineffective, the

material will not be removed from the tissue fluids or blood. Instead, it will stay in those fluids and increase in concentration. The failure of receptor-mediated endocytosis causes some human diseases. For example, receptor mediated endocytosis removes low density lipoprotein or LDL (or "bad" cholesterol) from the blood. In the human genetic disease familial hypercholesterolemia, the LDL receptors are defective or missing entirely. People with this condition have life-threatening levels of cholesterol in their blood, because their cells cannot clear LDL particles.

Although receptor-mediated endocytosis is designed to bring specific substances that are normally in the extracellular fluid into the cell, other substances may gain entry into the cell at the same site. Flu viruses, diphtheria, and cholera toxin all have sites that cross-react with normal receptor-binding sites and gain entry into cells.

Link to Learning

See receptor-mediated endocytosis in action, and click on different parts for a focused animation.

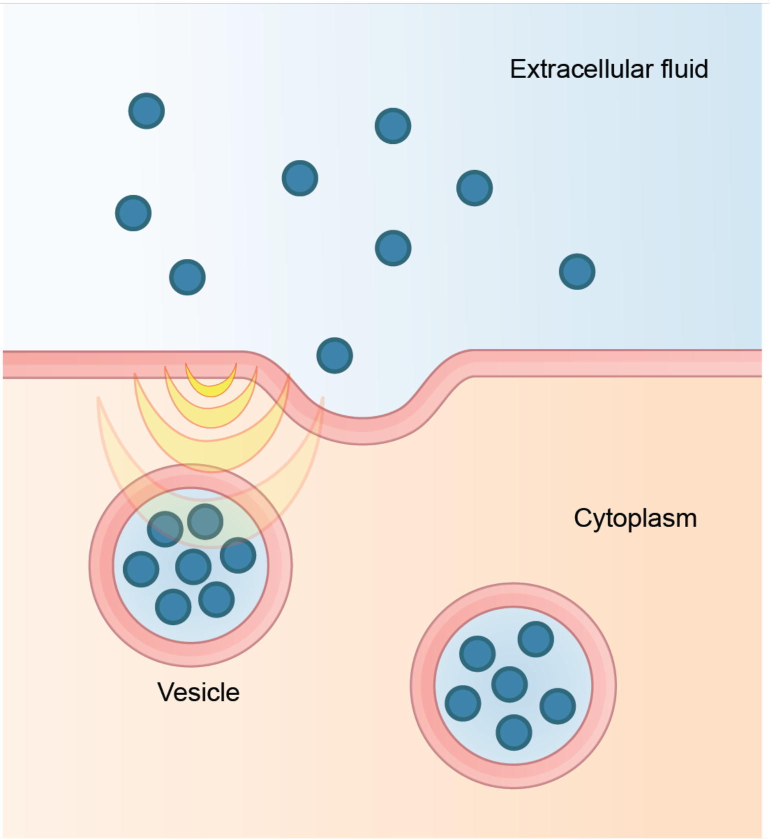
In exocytosis, vesicles containing substances fuse with the plasma membrane. The contents then

release to the cell's exterior. (credit: modification of work by Mariana Ruiz Villareal)

Exocytosis

The reverse process of moving material into a cell is the process of exocytosis. **Exocytosis** is the opposite of the processes we discussed above in that its purpose is to expel material from the cell into the extracellular fluid. Waste material is enveloped in a membrane and fuses with the plasma membrane's interior. This fusion opens the membranous envelope on the cell's exterior, and the waste material expels into the extracellular space ([link]). Other examples of cells releasing molecules via exocytosis include extracellular matrix protein secretion and neurotransmitter secretion into the synaptic cleft by synaptic vesicles.

Exocytosis



Methods of Transport, Energy Requirements, and Types of Transported

Material Transport Method	Active/Passive	Material Transported
Diffusion	Passive	Small-molecular weight material
Osmosis	Passive	Water
Facilitated transport/ diffusion	Passive	Sodium, potassium, calcium, glucose
Primary active transport	Active	Sodium, potassium, calcium
Secondary active transport	Active	Amino acids, lactose
Phagocytosis	Active	Large macromolecules, whole cells, or cellular structures
Pinocytosis and potocytosis	Active	Small molecules (liquids/water)
Receptor-mediated endocytosis	Active	Large quantities of macromolecules

Section Summary

Active transport methods require directly using ATP to fuel the transport. In a process scientists call

phagocytosis, other cells can engulf large particles, such as macromolecules, cell parts, or whole cells. In phagocytosis, a portion of the membrane invaginates and flows around the particle, eventually pinching off and leaving the particle entirely enclosed by a plasma membrane's envelope. The cell breaks down vesicle contents, with the particles either used as food or dispatched. Pinocytosis is a similar process on a smaller scale. The plasma membrane invaginates and pinches off, producing a small envelope of fluid from outside the cell. Pinocytosis imports substances that the cell needs from the extracellular fluid. The cell expels waste in a similar but reverse manner. It pushes a membranous vacuole to the plasma membrane, allowing the vacuole to fuse with the membrane and incorporate itself into the membrane structure, releasing its contents to the exterior.

Review Questions

What happens to the membrane of a vesicle after exocytosis?

1. It leaves the cell.
2. It is disassembled by the cell.
3. It fuses with and becomes part of the plasma membrane.

4. It is used again in another exocytosis event.

C

Which transport mechanism can bring whole cells into a cell?

1. pinocytosis
2. phagocytosis
3. facilitated transport
4. primary active transport

B

In what important way does receptor-mediated endocytosis differ from phagocytosis?

1. It transports only small amounts of fluid.
2. It does not involve the pinching off of membrane.
3. It brings in only a specifically targeted substance.
4. It brings substances into the cell, while phagocytosis removes substances.

C

Many viruses enter host cells through receptor-mediated endocytosis. What is an advantage of this entry strategy?

1. The virus directly enters the cytoplasm of the cell.
2. The virus is protected from recognition by white blood cells.
3. The virus only enters its target host cell type.
4. The virus can directly inject its genome into the cell’s nucleus.

C

Which of the following organelles relies on exocytosis to complete its function?

1. Golgi apparatus
2. vacuole
3. mitochondria
4. endoplasmic reticulum

A

Imagine a cell can perform exocytosis, but only minimal endocytosis. What would happen to the cell?

1. The cell would secrete all its intracellular proteins.
2. The plasma membrane would increase in size over time.
3. The cell would stop expressing integral receptor proteins in its plasma membrane.
4. The cell would lyse.

B

Critical Thinking Questions

Why is it important that there are different types of proteins in plasma membranes for the transport of materials into and out of a cell?

The proteins allow a cell to select what compound will be transported, meeting the needs of the cell and not bringing in anything else.

Why do ions have a difficult time getting through plasma membranes despite their small size?

Ions are charged, and consequently, they are hydrophilic and cannot associate with the lipid portion of the membrane. Ions must be transported by carrier proteins or ion channels.

Glossary

caveolin
: protein that coats the plasma membrane's cytoplasmic side and participates in the liquid uptake process by potocytosis

clathrin
: protein that coats the plasma membrane's inward-facing surface and assists in forming specialized structures, like coated pits, for phagocytosis

endocytosis
: type of active transport that moves substances, including fluids and particles, into a cell

exocytosis
: process of passing bulk material out of a cell

pinocytosis
: a variation of endocytosis that imports macromolecules that the cell needs from the extracellular fluid

potocytosis

variation of pinocytosis that uses a different coating protein (caveolin) on the plasma membrane's cytoplasmic side

receptor-mediated endocytosis

variation of endocytosis that involves using specific binding proteins in the plasma membrane for specific molecules or particles, and clathrin-coated pits that become clathrin-coated vesicles

Introduction

class="introduction" A hummingbird needs energy to maintain prolonged periods of flight. The bird obtains its energy from taking in food and transforming the nutrients into energy through a series of biochemical reactions. The flight muscles in birds are extremely efficient in energy production. (credit: modification of work by Cory Zanker)

Virtually every task performed by living organisms requires energy. Organisms require energy to perform heavy labor and exercise, but humans also use considerable energy while thinking, and even during sleep. Every organism's living cells constantly use energy. Organisms import nutrients and other molecules. They metabolize (break down) and possibly synthesize into new molecules. If necessary, molecules modify, move around the cell and may distribute themselves to the entire organism. For example, the large proteins that make up muscles are actively built from smaller molecules. Complex

carbohydrates break down into simple sugars that the cell uses for energy. Just as energy is required to both build and demolish a building, energy is required to synthesize and break down molecules. Additionally, signaling molecules such as hormones and neurotransmitters transport between cells. Cells ingest and break down bacteria and viruses. Cells must also export waste and toxins to stay healthy, and many cells must swim or move surrounding materials via the beating motion of cellular appendages like cilia and flagella.

The cellular processes that we listed above require a steady supply of energy. From where, and in what form, does this energy come? How do living cells obtain energy, and how do they use it? This chapter will discuss different forms of energy and the physical laws that govern energy transfer. This chapter will also describe how cells use energy and replenish it, and how chemical reactions in the cell perform with great efficiency.

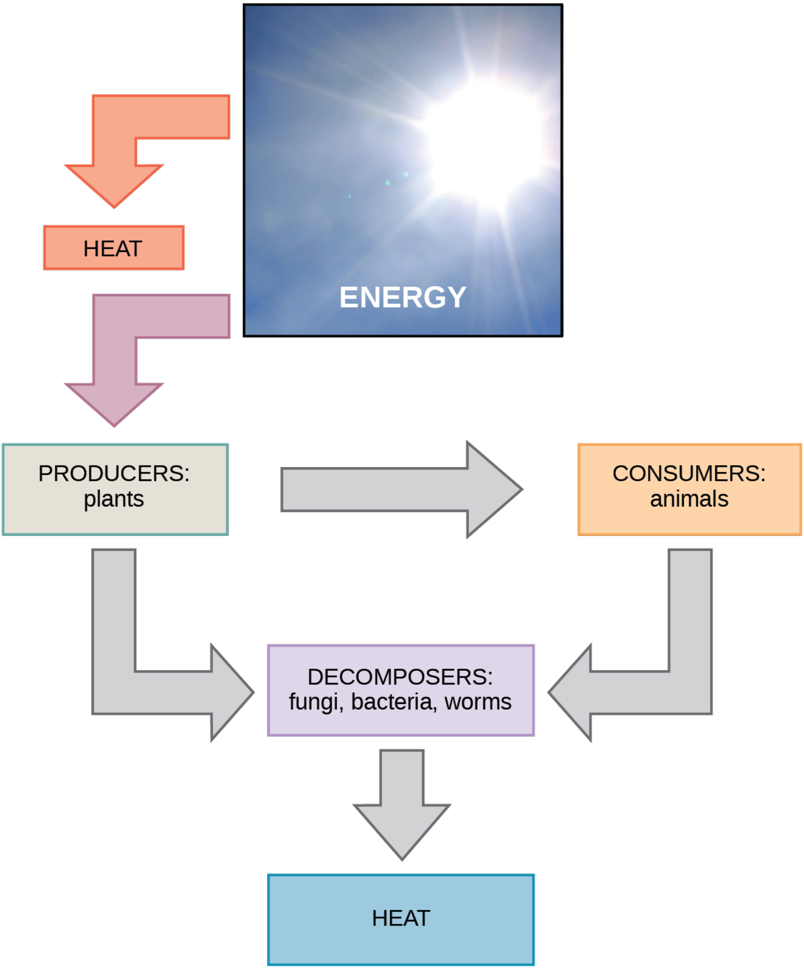
Energy and Metabolism

By the end of this section, you will be able to do the following:

- Explain metabolic pathways and describe the two major types
- Discuss how chemical reactions play a role in energy transfer

Scientists use the term **bioenergetics** to discuss the concept of energy flow ([link]) through living systems, such as cells. Cellular processes such as building and breaking down complex molecules occur through stepwise chemical reactions. Some of these chemical reactions are spontaneous and release energy; whereas, others require energy to proceed. Just as living things must continually consume food to replenish what they have used, cells must continually obtain more energy to replenish that which the many energy-requiring chemical reactions that constantly take place use. All of the chemical reactions that transpire inside cells, including those that use and release energy, are the cell’s **metabolism**.

Most life forms on earth obtain their energy from the sun. Plants use photosynthesis to capture sunlight, and herbivores eat those plants to obtain energy. Carnivores eat the herbivores, and decomposers digest plant and animal matter.



Plants, like this oak tree and acorn, use energy from sunlight to make sugar and other organic molecules. Both plants and animals (like this squirrel) use cellular respiration to derive energy from the organic molecules that plants originally produced. (credit “acorn”: modification of work by Noel Reynolds; credit “squirrel”: modification of work by Dawn Huczek)

Carbohydrate Metabolism

Sugar (chemical reactions) metabolism (a simple carbohydrate) is a classic example of the many cellular processes that use and produce energy. Living things consume sugar as a major energy source, because sugar molecules have considerable energy stored within their bonds. The following equation describes the breakdown of glucose, a simple sugar:

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O + \text{energy}$$

Consumed carbohydrates have their origins in photosynthesizing organisms like plants ([link]). During photosynthesis, plants use the energy of sunlight to convert carbon dioxide gas (CO_2) into sugar molecules, like glucose ($C_6H_{12}O_6$). Because this process involves synthesizing a larger, energy-storing molecule, it requires an energy input to proceed. The following equation (notice that it is the reverse of the previous equation) describes the synthesis of glucose:

$$6CO_2 + 6H_2O + \text{energy} \rightarrow C_6H_{12}O_6 + 6O_2$$

During photosynthesis chemical reactions, energy is in the form of a very high-energy molecule scientists call ATP, or adenosine triphosphate. This is the primary energy currency of all cells. Just as the dollar is the currency we use to buy goods, cells use ATP molecules as energy currency to perform immediate work. The sugar (glucose) is stored as

starch or glycogen. Energy-storing polymers like these break down into glucose to supply ATP molecules.

Solar energy is required to synthesize a glucose molecule during the photosynthesis reactions. In photosynthesis, light energy from the sun initially transforms into chemical energy that temporally stores itself in the energy carrier molecules ATP and NADPH (nicotinamide adenine dinucleotide phosphate). Photosynthesis later uses the stored energy in ATP and NADPH to build one glucose molecule from six molecules of CO_2. This process is analogous to eating breakfast in the morning to acquire energy for your body that you can use later in the day. Under ideal conditions, energy from 18 molecules of ATP is required to synthesize one glucose molecule during photosynthesis reactions. Glucose molecules can also combine with and convert into other sugar types. When an organism consumes sugars, glucose molecules eventually make their way into each organism's living cell. Inside the cell, each sugar molecule breaks down through a complex series of chemical reactions. The goal of these reactions is to harvest the energy stored inside the sugar molecules. The harvested energy makes high-energy ATP molecules, which perform work, powering many chemical reactions in the cell. The amount of energy needed to make one glucose molecule from six carbon dioxide molecules is 18 ATP molecules and 12 NADPH molecules (each

one of which is energetically equivalent to three ATP molecules), or a total of 54 molecule equivalents required for synthesizing one glucose molecule. This process is a fundamental and efficient way for cells to generate the molecular energy that they require.

Anabolic pathways are those that require energy to synthesize larger molecules. Catabolic pathways are those that generate energy by breaking down larger molecules. Both types of pathways are required for maintaining the cell's energy balance.

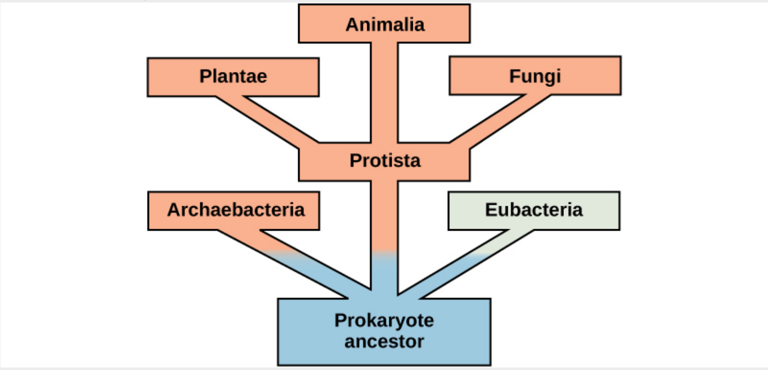
Metabolic Pathways

The processes of making and breaking down sugar molecules illustrate two types of metabolic pathways. A metabolic pathway is a series of interconnected biochemical reactions that convert a substrate molecule or molecules, step-by-step, through a series of metabolic intermediates, eventually yielding a final product or products. In the case of sugar metabolism, the first metabolic

pathway synthesized sugar from smaller molecules, and the other pathway broke sugar down into smaller molecules. Scientists call these two opposite processes—the first requiring energy and the second producing energy—anabolic (building) and catabolic (breaking down) pathways, respectively. Consequently, building (anabolism) and degradation (catabolism) comprise metabolism.

Evolution Connection

This tree shows the evolution of the various branches of life. The vertical dimension is time. Early life forms, in blue, used anaerobic metabolism to obtain energy from their surroundings.



Evolution of Metabolic Pathways

There is more to the complexity of metabolism than understanding the metabolic pathways alone. Metabolic complexity varies from organism to organism. Photosynthesis is the primary pathway in which photosynthetic organisms like plants

(planktonic algae perform the majority of global synthesis) harvest the sun's energy and convert it into carbohydrates. The by-product of photosynthesis is oxygen, which some cells require to carry out cellular respiration. During cellular respiration, oxygen aids in the catabolic breakdown of carbon compounds, like carbohydrates. Among the products are CO_2 and ATP. In addition, some eukaryotes perform catabolic processes without oxygen (fermentation); that is, they perform or use anaerobic metabolism. Organisms probably evolved anaerobic metabolism to survive (living organisms came into existence about 3.8 billion years ago, when the atmosphere lacked oxygen). Despite the differences between organisms and the complexity of metabolism, researchers have found that all branches of life share some of the same metabolic pathways, suggesting that all organisms evolved from the same ancient common ancestor ([link]). Evidence indicates that over time, the pathways diverged, adding specialized enzymes to allow organisms to better adapt to their environment, thus increasing their chance to survive. However, the underlying principle remains that all organisms must harvest energy from their environment and convert it to ATP to carry out cellular functions.

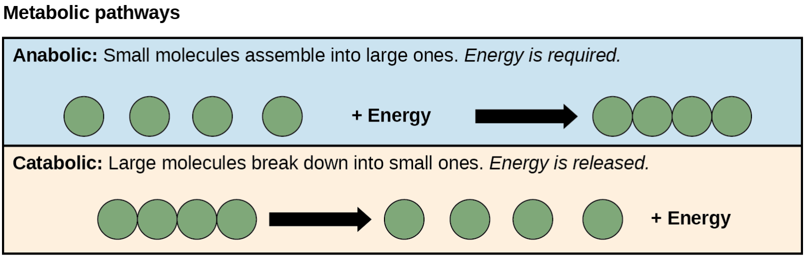
Anabolic and Catabolic Pathways

Anabolic pathways require an input of energy to synthesize complex molecules from simpler ones. Synthesizing sugar from CO_2 is one example. Other examples are synthesizing large proteins from amino acid building blocks, and synthesizing new DNA strands from nucleic acid building blocks. These biosynthetic processes are critical to the cell's life, take place constantly, and demand energy that ATP and other high-energy molecules like NADH (nicotinamide adenine dinucleotide) and NADPH provide ([link]).

ATP is an important molecule for cells to have in sufficient supply at all times. The breakdown of sugars illustrates how a single glucose molecule can store enough energy to make a great deal of ATP, 36 to 38 molecules. This is a **catabolic** pathway. Catabolic pathways involve degrading (or breaking down) complex molecules into simpler ones. Molecular energy stored in complex molecule bonds release in catabolic pathways and harvest in such a way that it can produce ATP. Other energy-storing molecules, such as fats, also break down through similar catabolic reactions to release energy and make ATP ([link]).

It is important to know that metabolic pathway chemical reactions do not take place spontaneously. A protein called an enzyme facilitates or catalyzes each reaction step. Enzymes are important for catalyzing all types of biological reactions—those

that require energy as well as those that release energy.



Section Summary

Cells perform the functions of life through various chemical reactions. A cell's metabolism refers to the chemical reactions that take place within it. There are metabolic reactions that involve breaking down complex chemicals into simpler ones, such as breaking down large macromolecules. Scientists refer to this process as catabolism, and we associate such reactions an energy release. On the other end of the spectrum, anabolism refers to metabolic processes that build complex molecules out of simpler ones, such as macromolecule synthesis. Anabolic processes require energy. Glucose synthesis and glucose breakdown are examples of anabolic and catabolic pathways, respectively.

Multiple Choice

Energy is stored long-term in the bonds of ___ and used short-term to perform work from a(n) ___ molecule.

1. ATP : glucose
2. an anabolic molecule : catabolic molecule
3. glucose : ATP
4. a catabolic molecule : anabolic molecule

C

DNA replication involves unwinding two strands of parent DNA, copying each strand to synthesize complementary strands, and releasing the parent and daughter DNA. Which of the following accurately describes this process?

1. This is an anabolic process.
2. This is a catabolic process.
3. This is both anabolic and catabolic.
4. This is a metabolic process but is neither anabolic nor catabolic.

A

Critical Thinking Questions

Does physical exercise involve anabolic and/or catabolic processes? Give evidence for your answer.

Physical exercise involves both anabolic and catabolic processes. Body cells break down sugars to provide ATP to do the work necessary for exercise, such as muscle contractions. This is catabolism. Muscle cells also must repair muscle tissue damaged by exercise by building new muscle. This is anabolism.

Name two different cellular functions that require energy that parallel human energy-requiring functions.

Energy is required for cellular motion, through beating of cilia or flagella, as well as human motion, produced by muscle contraction. Cells also need energy to perform digestion, as humans require energy to digest food.

Glossary

anabolic
: (also, anabolism) pathways that require an

energy input to synthesize complex molecules from simpler ones

bioenergetics
: study of energy flowing through living systems

catabolic
: (also, catabolism) pathways in which complex molecules break down into simpler ones

metabolism
: all the chemical reactions that take place inside cells, including anabolism and catabolism

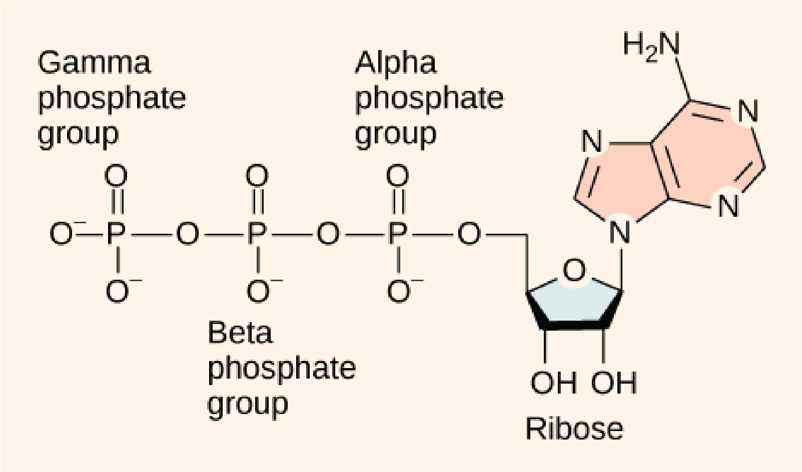
ATP: Adenosine Triphosphate

By the end of this section, you will be able to do the following:

- Explain ATP's role as the cellular energy currency
- Describe how energy releases through ATP hydrolysis

Even exergonic, energy-releasing reactions require a small amount of activation energy in order to proceed. However, consider endergonic reactions, which require much more energy input, because their products have more free energy than their reactants. Within the cell, from where does energy to power such reactions come? The answer lies with an energy-supplying molecule scientists call **adenosine triphosphate**, or **ATP**. This is a small, relatively simple molecule ([link]), but within some of its bonds, it contains the potential for a quick burst of energy that can be harnessed to perform cellular work. Think of this molecule as the cells' primary energy currency in much the same way that money is the currency that people exchange for things they need. ATP powers the majority of energy-requiring cellular reactions.

ATP is the cell's primary energy currency. It has an adenosine backbone with three phosphate groups attached.



As its name suggests, adenosine triphosphate is comprised of adenosine bound to three phosphate groups ([link]). Adenosine is a nucleoside consisting of the nitrogenous base adenine and a five-carbon sugar, ribose. The three phosphate groups, in order of closest to furthest from the ribose sugar, are alpha, beta, and gamma. Together, these chemical groups constitute an energy powerhouse. However, not all bonds within this molecule exist in a particularly high-energy state. Both bonds that link the phosphates are equally high-energy bonds (**phosphoanhydride bonds**) that, when broken, release sufficient energy to power a variety of cellular reactions and processes. These high-energy bonds are the bonds between the second and third (or beta and gamma) phosphate groups and between the first and second phosphate groups. These bonds are "high-energy" because the products of such bond breaking—adenosine diphosphate (ADP) and one inorganic phosphate group (P_i)—have

considerably lower free energy than the reactants: ATP and a water molecule. Because this reaction takes place using a water molecule, it is a hydrolysis reaction. In other words, ATP hydrolyzes into ADP in the following reaction:

$$\text{ATP} + \text{H}_2\text{O} \rightarrow \text{ADP} + \text{P}_i + \text{free energy}$$

Like most chemical reactions, ATP to ADP hydrolysis is reversible. The reverse reaction regenerates ATP from ADP + P_i. Cells rely on ATP regeneration just as people rely on regenerating spent money through some sort of income. Since ATP hydrolysis releases energy, ATP regeneration must require an input of free energy. This equation expresses ATP formation:

$$\text{ADP} + \text{P}_i + \text{free energy} \rightarrow \text{ATP} + \text{H}_2\text{O}$$

Two prominent questions remain with regard to using ATP as an energy source. Exactly how much free energy releases with ATP hydrolysis, and how does that free energy do cellular work? The calculated ∆G for the hydrolysis of one ATP mole into ADP and P_i is -7.3 kcal/mole (-30.5 kJ/mol). Since this calculation is true under standard conditions, one would expect a different value exists under cellular conditions. In fact, the ∆G for one ATP mole's hydrolysis in a living cell is almost double the value at standard conditions: –14 kcal/mol (-57 kJ/mol).

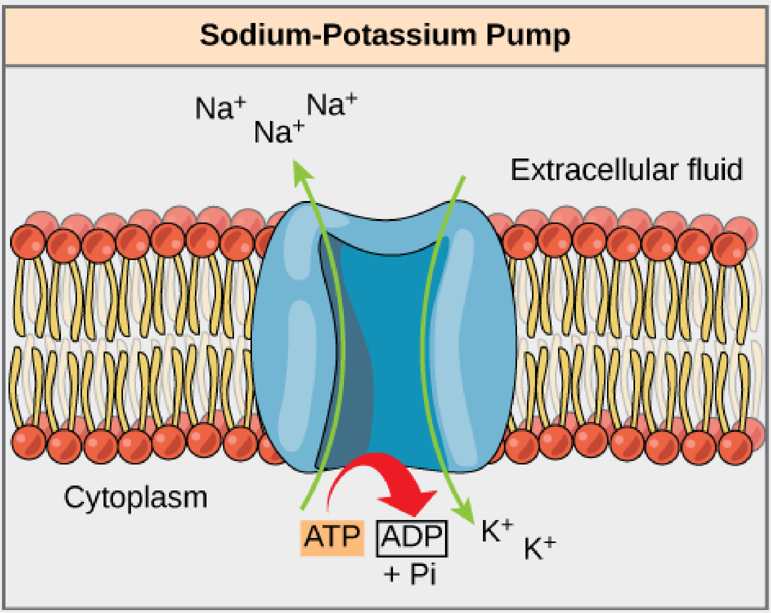
ATP is a highly unstable molecule. Unless quickly

used to perform work, ATP spontaneously dissociates into ADP + P_i, and the free energy released during this process is lost as heat. The second question we posed above discusses how ATP hydrolysis energy release performs work inside the cell. This depends on a strategy scientists call energy coupling. Cells couple the ATP hydrolysis' exergonic reaction allowing them to proceed. One example of energy coupling using ATP involves a transmembrane ion pump that is extremely important for cellular function. This sodium-potassium pump (Na^+/K^+ pump) drives sodium out of the cell and potassium into the cell ([link]). A large percentage of a cell's ATP powers this pump, because cellular processes bring considerable sodium into the cell and potassium out of it. The pump works constantly to stabilize cellular concentrations of sodium and potassium. In order for the pump to turn one cycle (exporting three Na^+ ions and importing two K^+ ions), one ATP molecule must hydrolyze. When ATP hydrolyzes, its gamma phosphate does not simply float away, but it actually transfers onto the pump protein. Scientists call this process of a phosphate group binding to a molecule phosphorylation. As with most ATP hydrolysis cases, a phosphate from ATP transfers onto another molecule. In a phosphorylated state, the Na^+/K^+ pump has more free energy and is triggered to undergo a conformational change. This change allows it to release Na^+ to the cell's outside. It then binds extracellular K^+, which, through

another conformational change, causes the phosphate to detach from the pump. This phosphate release triggers the K^+ to release to the cell's inside. Essentially, the energy released from the ATP hydrolysis couples with the energy required to power the pump and transport Na^+ and K^+ ions. ATP performs cellular work using this basic form of energy coupling through phosphorylation.

Visual Connection

The sodium-potassium pump is an example of energy coupling. The energy derived from exergonic ATP hydrolysis pumps sodium and potassium ions across the cell membrane.



One ATP molecule's hydrolysis releases 7.3 kcal/mol of energy ($\Delta G = -7.3$ kcal/mol of energy). If

it takes 2.1 kcal/mol of energy to move one $Na+$ across the membrane ($\Delta G = +2.1$ kcal/mol of energy), how many sodium ions could one ATP molecule's hydrolysis move?

Often during cellular metabolic reactions, such as nutrient synthesis and breakdown, certain molecules must alter slightly in their conformation to become substrates for the next step in the reaction series. One example is during the very first steps of cellular respiration, when a sugar glucose molecule breaks down in the process of glycolysis. In the first step, ATP is required to phosphorylze glucose, creating a high-energy but unstable intermediate. This phosphorylation reaction powers a conformational change that allows the phosphorylated glucose molecule to convert to the phosphorylated sugar fructose. Fructose is a necessary intermediate for glycolysis to move forward. Here, ATP hydrolysis' exergonic reaction couples with the endergonic reaction of converting glucose into a phosphorylated intermediate in the pathway. Once again, the energy released by breaking a phosphate bond within ATP was used for phosphorylyzing another molecule, creating an unstable intermediate and powering an important conformational change.

Link to Learning

See an interactive animation of the ATP-producing glycolysis process at this site.

Section Summary

ATP is the primary energy-supplying molecule for living cells. ATP is comprised of a nucleotide, a five-carbon sugar, and three phosphate groups. The bonds that connect the phosphates (phosphoanhydride bonds) have high-energy content. The energy released from ATP hydrolysis into ADP + P_i performs cellular work. Cells use ATP to perform work by coupling ATP hydrolysis' exergonic reaction with endergonic reactions. ATP donates its phosphate group to another molecule via phosphorylation. The phosphorylated molecule is at a higher-energy state and is less stable than its unphosphorylated form, and this added energy from phosphate allows the molecule to undergo its endergonic reaction.

Visual Connection Questions

[link] One ATP molecule's hydrolysis releases 7.3 kcal/mol of energy ($\Delta G = -7.3$ kcal/mol of energy). If it takes 2.1 kcal/mol of energy to move one $Na+$ across the membrane ($\Delta G = +2.1$ kcal/mol of energy), how many sodium ions could one ATP molecule's hydrolysis move?

[link] Three sodium ions could be moved by the hydrolysis of one ATP molecule. The ΔG of the coupled reaction must be negative. Movement of three sodium ions across the membrane will take 6.3 kcal of energy (2.1 kcal $\times$ 3 $Na+$ ions $=$ 6.3 kcal). Hydrolysis of ATP provides 7.3 kcal of energy, more than enough to power this reaction. Movement of four sodium ions across the membrane, however, would require 8.4 kcal of energy, more than one ATP molecule can provide.

Review Questions

The energy released by the hydrolysis of ATP is___

1. primarily stored between the alpha and beta phosphates

2. equal to -57 kcal/mol
3. harnessed as heat energy by the cell to perform work
4. providing energy to coupled reactions

D

Which of the following molecules is likely to have the most potential energy?

1. sucrose
2. ATP
3. glucose
4. ADP

A

Critical Thinking Questions

Do you think that the E_A for ATP hydrolysis is relatively low or high? Explain your reasoning.

The activation energy for hydrolysis is very low. Not only is ATP hydrolysis an exergonic process with a large $-\Delta G$, but ATP is also a

very unstable molecule that rapidly breaks down into ADP + P_i if not utilized quickly. This suggests a very low E_A since it hydrolyzes so quickly.

Glossary

ATP
: adenosine triphosphate, the cell's energy currency

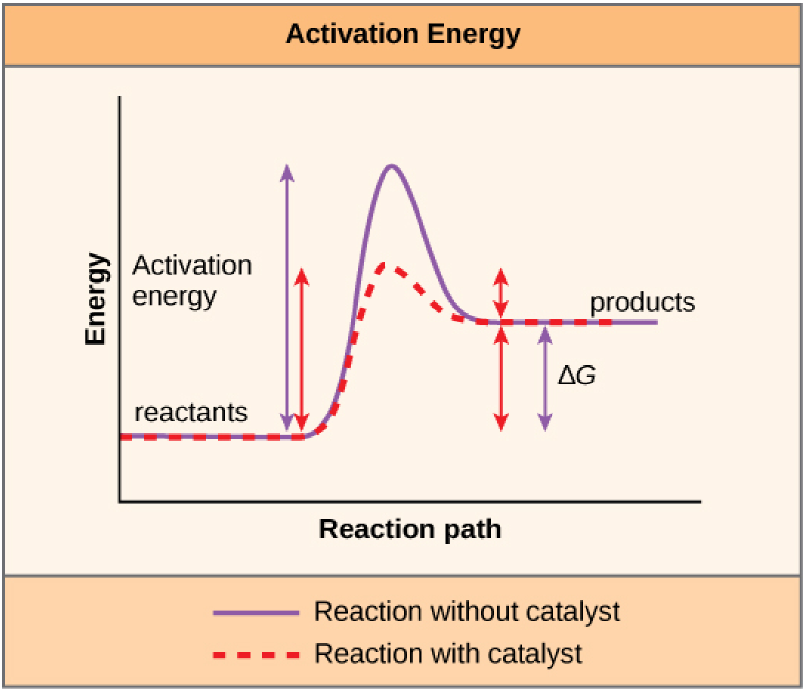
phosphoanhydride bond
: bond that connects phosphates in an ATP molecule

Enzymes

By the end of this section, you will be able to do the following:

- Describe the role of enzymes in metabolic pathways
- Explain how enzymes function as molecular catalysts
- Discuss enzyme regulation by various factors

A substance that helps a chemical reaction to occur is a catalyst, and the special molecules that catalyze biochemical reactions are enzymes. Almost all enzymes are proteins, comprised of amino acid chains, and they perform the critical task of lowering the activation energies of chemical reactions inside the cell. Enzymes do this by binding to the reactant molecules, and holding them in such a way as to make the chemical bond-breaking and bond-forming processes take place more readily. It is important to remember that enzymes do not change the reaction's ΔG. In other words, they do not change whether a reaction is exergonic (spontaneous) or endergonic. This is because they do not change the reactants' or products' free energy. They only reduce the activation energy required to reach the transition state ([link]). Enzymes lower the reaction's activation energy but do not change the reaction's free energy.



According to the induced-fit model, both enzyme and substrate undergo dynamic conformational changes upon binding. The enzyme contorts the substrate into its transition state, thereby increasing the reaction's rate.

Enzyme Active Site and Substrate Specificity

The chemical reactants to which an enzyme binds are the enzyme's **substrates**. There may be one or more substrates, depending on the particular chemical reaction. In some reactions, a single-reactant substrate breaks down into multiple products. In others, two substrates may come

together to create one larger molecule. Two reactants might also enter a reaction, both become modified, and leave the reaction as two products. The location within the enzyme where the substrate binds is the enzyme's **active site**. This is where the "action" happens. Since enzymes are proteins, there is a unique combination of amino acid residues (also side chains, or R groups) within the active site. Different properties characterize each residue. These can be large or small, weakly acidic or basic, hydrophilic or hydrophobic, positively or negatively charged, or neutral. The unique combination of amino acid residues, their positions, sequences, structures, and properties, creates a very specific chemical environment within the active site. This specific environment is suited to bind, albeit briefly, to a specific chemical substrate (or substrates). Due to this jigsaw puzzle-like match between an enzyme and its substrates (which adapts to find the best fit between the transition state and the active site), enzymes are known for their specificity. The "best fit" results from the shape and the amino acid functional group's attraction to the substrate. There is a specifically matched enzyme for each substrate and, thus, for each chemical reaction; however, there is flexibility as well.

The fact that active sites are so perfectly suited to provide specific environmental conditions also means that they are subject to local enviromental influences. It is true that increasing the

environmental temperature generally increases reaction rates, enzyme-catalyzed or otherwise. However, increasing or decreasing the temperature outside of an optimal range can affect chemical bonds within the active site in such a way that they are less well suited to bind substrates. High temperatures will eventually cause enzymes, like other biological molecules, to **denature**, a process that changes the substance's natural properties. Likewise, the local environment's pH can also affect enzyme function. Active site amino acid residues have their own acidic or basic properties that are optimal for catalysis. These residues are sensitive to changes in pH that can impair the way substrate molecules bind. Enzymes are suited to function best within a certain pH range, and, as with temperature, extreme environmental pH values (acidic or basic) can cause enzymes to denature.

Induced Fit and Enzyme Function

For many years, scientists thought that enzyme-substrate binding took place in a simple "lock-and-key" fashion. This model asserted that the enzyme and substrate fit together perfectly in one instantaneous step. However, current research supports a more refined view scientists call **induced fit** ([link]). This model expands upon the lock-and-key model by describing a more dynamic interaction between enzyme and substrate. As the enzyme and substrate come together, their interaction causes a

mild shift in the enzyme’s structure that confirms an ideal binding arrangement between the enzyme and the substrate's transition state. This ideal binding maximizes the enzyme’s ability to catalyze its reaction.

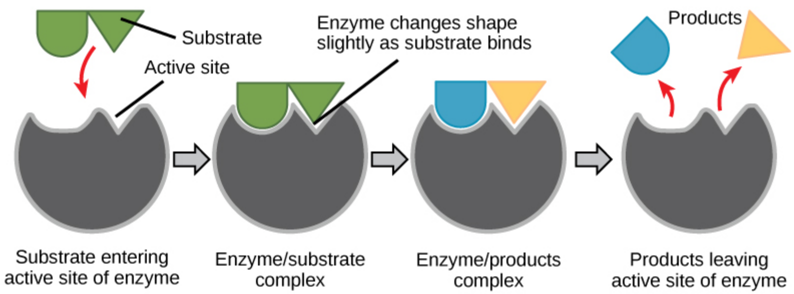
Link to Learning

View an induced fit animation at this website.

When an enzyme binds its substrate, it forms an enzyme-substrate complex. This complex lowers the reaction's activation energy and promotes its rapid progression in one of many ways. On a basic level, enzymes promote chemical reactions that involve more than one substrate by bringing the substrates together in an optimal orientation. The appropriate region (atoms and bonds) of one molecule is juxtaposed to the other molecule's appropriate region with which it must react. Another way in which enzymes promote substrate reaction is by creating an optimal environment within the active site for the reaction to occur. Certain chemical reactions might proceed best in a slightly acidic or non-polar environment. The chemical properties that emerge from the particular arrangement of amino acid residues within an active site create the perfect environment for an enzyme’s specific

substrates to react.

You have learned that the activation energy required for many reactions includes the energy involved in manipulating or slightly contorting chemical bonds so that they can easily break and allow others to reform. Enzymatic action can aid this process. The enzyme-substrate complex can lower the activation energy by contorting substrate molecules in such a way as to facilitate bond-breaking, helping to reach the transition state. Finally, enzymes can also lower activation energies by taking part in the chemical reaction itself. The amino acid residues can provide certain ions or chemical groups that actually form covalent bonds with substrate molecules as a necessary step of the reaction process. In these cases, it is important to remember that the enzyme will always return to its original state at the reaction's completion. One of enzymes' hallmark properties is that they remain ultimately unchanged by the reactions they catalyze. After an enzyme catalyzes a reaction, it releases its product(s).



Competitive and noncompetitive inhibition affect

the reaction's rate differently. Competitive inhibitors affect the initial rate but do not affect the maximal rate; whereas, noncompetitive inhibitors affect the maximal rate. Allosteric inhibitors modify the enzyme's active site so that substrate binding is reduced or prevented. In contrast, allosteric activators modify the enzyme's active site so that the affinity for the substrate increases. Vitamins are important coenzymes or precursors of coenzymes, and are required for enzymes to function properly. Multivitamin capsules usually contain mixtures of all the vitamins at different percentages. Metabolic pathways are a series of reactions that multiple enzymes catalyze. Feedback inhibition, where the pathway's end product inhibits an upstream step, is an important regulatory mechanism in cells.

Metabolism Control Through Enzyme Regulation

It would seem ideal to have a scenario in which all the encoded enzymes in an organism’s genome existed in abundant supply and functioned optimally under all cellular conditions, in all cells, at all times. In reality, this is far from the case. A variety of mechanisms ensure that this does not happen. Cellular needs and conditions vary from cell to cell, and change within individual cells over time. The required enzymes and energetic demands of stomach cells are different from those of fat storage cells, skin cells, blood cells, and nerve cells.

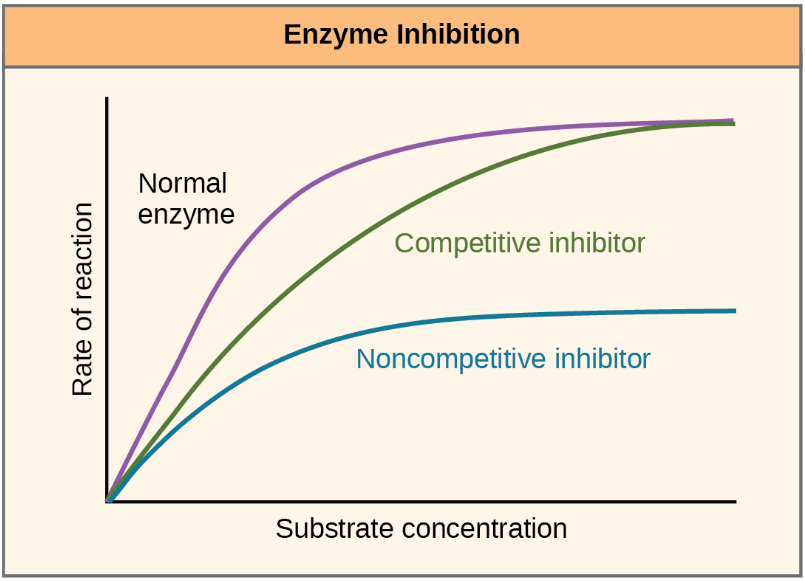
Furthermore, a digestive cell works much harder to process and break down nutrients during the time that closely follows a meal compared with many hours after a meal. As these cellular demands and conditions vary, so do the amounts and functionality of different enzymes.

Since the rates of biochemical reactions are controlled by activation energy, and enzymes lower and determine activation energies for chemical reactions, the relative amounts and functioning of the variety of enzymes within a cell ultimately determine which reactions will proceed and at which rates. This determination is tightly controlled. In certain cellular environments, environmental factors like pH and temperature partly control enzyme activity. There are other mechanisms through which cells control enzyme activity and determine the rates at which various biochemical reactions will occur.

Molecular Regulation of Enzymes

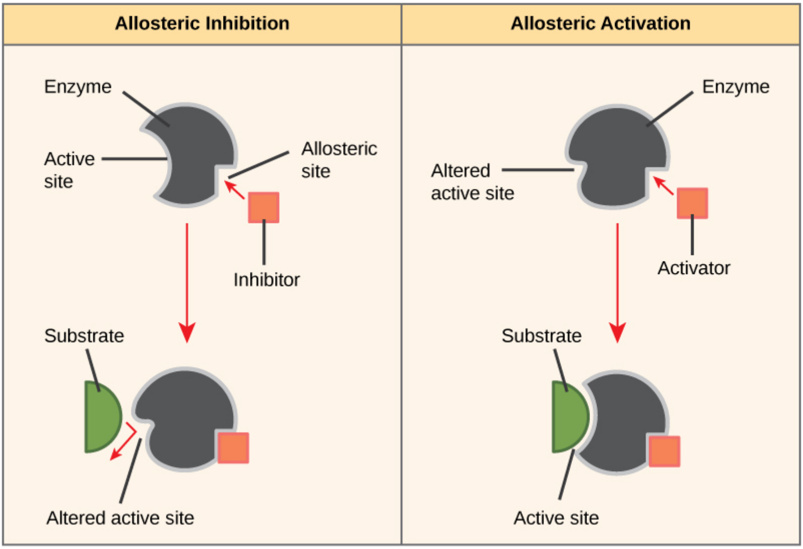
Enzymes can be regulated in ways that either promote or reduce their activity. There are many different kinds of molecules that inhibit or promote enzyme function, and various mechanisms exist for doing so. For example, in some cases of enzyme inhibition, an inhibitor molecule is similar enough to a substrate that it can bind to the active site and simply block the substrate from binding. When this

happens, the enzyme is inhibited through **competitive inhibition**, because an inhibitor molecule competes with the substrate for active site binding ([link]). Alternatively, in noncompetitive inhibition, an inhibitor molecule binds to the enzyme in a location other than an allosteric site, a binding site away from the active site, and still manages to block substrate binding to the active site.



Some inhibitor molecules bind to enzymes in a location where their binding induces a conformational change that reduces the enzyme's affinity for its substrate. This type of inhibition is an **allosteric inhibition** ([link]). More than one polypeptide comprise most allosterically regulated enzymes, meaning that they have more than one protein subunit. When an allosteric inhibitor binds to an enzyme, all active sites on the protein subunits

change slightly such that they bind their substrates with less efficiency. There are allosteric activators as well as inhibitors. Allosteric activators bind to locations on an enzyme away from the active site, inducing a conformational change that increases the affinity of the enzyme’s active site(s) for its substrate(s).



Everyday Connection

Have you ever wondered how pharmaceutical drugs are developed? (credit: Deborah Austin)

Drug Discovery by Looking for Inhibitors of Key Enzymes in Specific Pathways

Enzymes are key components of metabolic pathways. Understanding how enzymes work and how they can be regulated is a key principle behind developing many pharmaceutical drugs ([link]) on the market today. Biologists working in this field collaborate with other scientists, usually chemists, to design drugs.

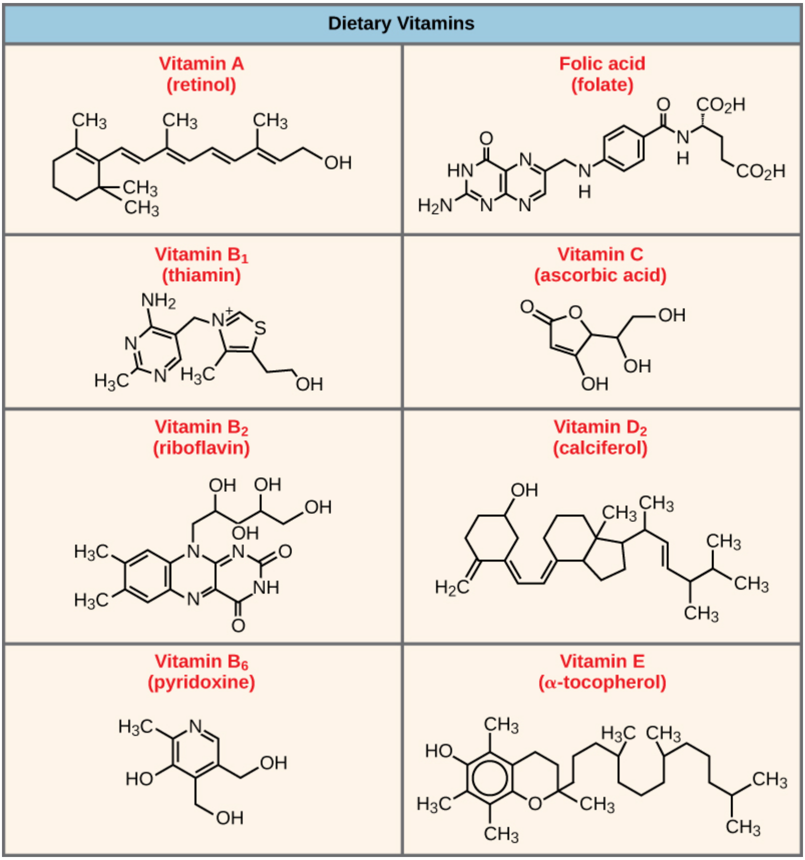
Consider statins for example—which is a class of drugs that reduces cholesterol levels. These compounds are essentially inhibitors of the enzyme HMG-CoA reductase. HMG-CoA reductase is the enzyme that synthesizes cholesterol from lipids in the body. By inhibiting this enzyme, the drug reduces cholesterol levels synthesized in the body. Similarly, acetaminophen, popularly marketed under the brand name Tylenol, is an inhibitor of

the enzyme cyclooxygenase. While it is effective in providing relief from fever and inflammation (pain), scientists still do not completely understand its mechanism of action.

How are drugs developed? One of the first challenges in drug development is identifying the specific molecule that the drug is intended to target. In the case of statins, HMG-CoA reductase is the drug target. Researchers identify targets through painstaking research in the laboratory. Identifying the target alone is not sufficient. Scientists also need to know how the target acts inside the cell and which reactions go awry in the case of disease. Once researchers identify the target and the pathway, then the actual drug design process begins. During this stage, chemists and biologists work together to design and synthesize molecules that can either block or activate a particular reaction. However, this is only the beginning: both if and when a drug prototype is successful in performing its function, then it must undergo many tests from *in vitro* experiments to clinical trials before it can obtain FDA approval to be on the market.

Many enzymes don’t work optimally, or even at all, unless bound to other specific non-protein helper molecules, either temporarily through ionic or hydrogen bonds or permanently through stronger

covalent bonds. Two types of helper molecules are **cofactors** and **coenzymes**. Binding to these molecules promotes optimal conformation and function for their respective enzymes. Cofactors are inorganic ions such as iron (Fe^{++}) and magnesium (Mg^{++}). One example of an enzyme that requires a metal ion as a cofactor is the enzyme that builds DNA molecules, DNA polymerase, which requires a bound zinc ion (Zn^{++}) to function. Coenzymes are organic helper molecules, with a basic atomic structure comprised of carbon and hydrogen, which are required for enzyme action. The most common sources of coenzymes are dietary vitamins ([link]). Some vitamins are precursors to coenzymes and others act directly as coenzymes. Vitamin C is a coenzyme for multiple enzymes that take part in building the important connective tissue component, collagen. An important step in breaking down glucose to yield energy is catalysis by a multi-enzyme complex scientists call pyruvate dehydrogenase. Pyruvate dehydrogenase is a complex of several enzymes that actually requires one cofactor (a magnesium ion) and five different organic coenzymes to catalyze its specific chemical reaction. Therefore, enzyme function is, in part, regulated by an abundance of various cofactors and coenzymes, which the diets of most organisms supply.



Enzyme Compartmentalization

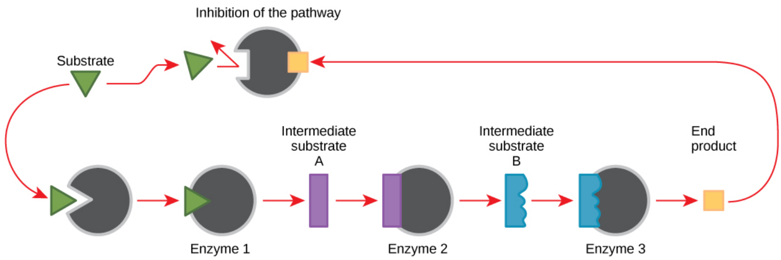
In eukaryotic cells, molecules such as enzymes are usually compartmentalized into different organelles. This allows for yet another level of regulation of enzyme activity. Enzymes required only for certain cellular processes are sometimes housed separately along with their substrates, allowing for more efficient chemical reactions. Examples of this sort of enzyme regulation based on location and proximity include the enzymes involved in the latter stages of

cellular respiration, which take place exclusively in the mitochondria, and the enzymes involved in digesting cellular debris and foreign materials, located within lysosomes.

Feedback Inhibition in Metabolic Pathways

Molecules can regulate enzyme function in many ways. However, a major question remains: What are these molecules and from where do they come? Some are cofactors and coenzymes, ions, and organic molecules, as you have learned. What other molecules in the cell provide enzymatic regulation, such as allosteric modulation, and competitive and noncompetitive inhibition? The answer is that a wide variety of molecules can perform these roles. Some include pharmaceutical and non-pharmaceutical drugs, toxins, and poisons from the environment. Perhaps the most relevant sources of enzyme regulatory molecules, with respect to cellular metabolism, are cellular metabolic reaction products themselves. In a most efficient and elegant way, cells have evolved to use their own reactions' products for feedback inhibition of enzyme activity. **Feedback inhibition** involves using a reaction product to regulate its own further production ([link]). The cell responds to the abundance of specific products by slowing down production during anabolic or catabolic reactions. Such reaction products may inhibit the enzymes that catalyzed their production through the mechanisms that we

described above.



Producing both amino acids and nucleotides is controlled through feedback inhibition. Additionally, ATP is an allosteric regulator of some of the enzymes involved in sugar's catabolic breakdown, the process that produces ATP. In this way, when ATP is abundant, the cell can prevent its further production. Remember that ATP is an unstable molecule that can spontaneously dissociate into ADP. If too much ATP were present in a cell, much of it would go to waste. Alternatively, ADP serves as a positive allosteric regulator (an allosteric activator) for some of the same enzymes that ATP inhibits. Thus, when relative ADP levels are high compared to ATP, the cell is triggered to produce more ATP through sugar catabolism.

Section Summary

Enzymes are chemical catalysts that accelerate chemical reactions at physiological temperatures by lowering their activation energy. Enzymes are usually proteins consisting of one or more

polypeptide chains. Enzymes have an active site that provides a unique chemical environment, comprised of certain amino acid R groups (residues). This unique environment is perfectly suited to convert particular chemical reactants for that enzyme, scientists call substrates, into unstable intermediates that they call transition states. Enzymes and substrates bind with an induced fit, which means that enzymes undergo slight conformational adjustments upon substrate contact, leading to full, optimal binding. Enzymes bind to substrates and catalyze reactions in four different ways: bringing substrates together in an optimal orientation, compromising the bond structures of substrates so that bonds can break down more easily, providing optimal environmental conditions for a reaction to occur, or participating directly in their chemical reaction by forming transient covalent bonds with the substrates.

Enzyme action must be regulated so that in a given cell at a given time, the desired reactions catalyze and the undesired reactions are not. Enzymes are regulated by cellular conditions, such as temperature and pH. They are also regulated through their location within a cell, sometimes compartmentalized so that they can only catalyze reactions under certain circumstances. Enzyme inhibition and activation via other molecules are other important ways that enzymes are regulated. Inhibitors can act competitively, noncompetitively,

or allosterically. Noncompetitive inhibitors are usually allosteric. Activators can also enhance enzyme function allosterically. The most common method by which cells regulate the enzymes in metabolic pathways is through feedback inhibition. During feedback inhibition, metabolic pathway products serve as inhibitors (usually allosteric) of one or more of the enzymes (usually the first committed enzyme of the pathway) involved in the pathway that produces them.

Review Questions

Which of the following is not true about enzymes:

1. They increase ΔG of reactions.
2. They are usually made of amino acids.
3. They lower the activation energy of chemical reactions.
4. Each one is specific to the particular substrate(s) to which it binds.

A

An allosteric inhibitor does which of the following?

1. Binds to an enzyme away from the active site and changes the conformation of the active site, increasing its affinity for substrate binding.
2. Binds to the active site and blocks it from binding substrate.
3. Binds to an enzyme away from the active site and changes the conformation of the active site, decreasing its affinity for the substrate.
4. Binds directly to the active site and mimics the substrate.

C

Which of the following analogies best describes the induced-fit model of enzyme-substrate binding?

1. a hug between two people
2. a key fitting into a lock
3. a square peg fitting through the square hole and a round peg fitting through the round hole of a children's toy
4. the fitting together of two jigsaw puzzle pieces

A

Critical Thinking Questions

With regard to enzymes, why are vitamins necessary for good health? Give examples.

Most vitamins and minerals act as coenzymes and cofactors for enzyme action. Many enzymes require the binding of certain cofactors or coenzymes to be able to catalyze their reactions. Since enzymes catalyze many important reactions, it is critical to obtain sufficient vitamins and minerals from the diet and from supplements. Vitamin C (ascorbic acid) is a coenzyme necessary for the action of enzymes that build collagen, an important protein component of connective tissue throughout the body. Magnesium ion (Mg^{++}) is an important cofactor that is necessary for the enzyme pyruvate dehydrogenase to catalyze part of the pathway that breaks down sugar to produce energy. Vitamins cannot be produced in the human body and therefore must be obtained in the diet.

Explain in your own words how enzyme feedback inhibition benefits a cell.

Feedback inhibition allows cells to control the amounts of metabolic products produced. If there is too much of a particular product relative to the cell’s needs, feedback inhibition effectively causes the cell to decrease production of that particular product. In general, this reduces the production of superfluous products and conserves energy, maximizing energy efficiency.

Glossary

active site
: enzyme's specific region to which the substrate binds

allosteric inhibition
: inhibition by a binding event at a site different from the active site, which induces a conformational change and reduces the enzyme's affinity for its substrate

coenzyme
: small organic molecule, such as a vitamin or its derivative, which is required to enhance an enzyme's activity

cofactor
: inorganic ion, such as iron and magnesium ions, required for optimal enzyme activity

regulation

competitive inhibition
: type of inhibition in which the inhibitor competes with the substrate molecule by binding to the enzyme's active site

denature
: process that changes a subtance's natural properties

feedback inhibition
: a product's effect of a reaction sequence to decrease its further production by inhibiting the first enzyme's activity in the pathway that produces it

induced fit
: dynamic fit between the enzyme and its substrate, in which both components modify their structures to allow for ideal binding

substrate
: molecule on which the enzyme acts

Introduction

class="introduction" This geothermal energy plant transforms thermal energy from deep in the ground into electrical energy, which can be easily used. (credit: modification of work by the U.S. Department of Defense)

The electrical energy plant in [link] converts energy from one form to another form that can be more easily used. This type of generating plant starts with underground thermal energy (heat) and transforms it into electrical energy that will be transported to homes and factories. Like a generating plant, plants and animals also must take in energy from the environment and convert it into a form that their cells can use. Mass and its stored energy enter an organism's body in one form and are converted into another form that can fuel the organism's life functions. In the process of photosynthesis, plants and other photosynthetic producers take in energy in the form of light (solar energy) and convert it into chemical energy in the form of glucose, which

stores this energy in its chemical bonds. Then, a series of metabolic pathways, collectively called cellular respiration, extracts the energy from the bonds in glucose and converts it into a form that all living things can use.

Energy in Living Systems
By the end of this section, you will be able to do the following:

- Discuss the importance of electrons in the transfer of energy in living systems
- Explain how ATP is used by cells as an energy source

Energy production within a cell involves many coordinated chemical pathways. Most of these pathways are combinations of oxidation and reduction reactions, which occur at the same time. An oxidation reaction strips an electron from an atom in a compound, and the addition of this electron to another compound is a reduction reaction. Because oxidation and reduction usually occur together, these pairs of reactions are called oxidation reduction reactions, or **redox reactions**. The oxidized form of the electron carrier (NAD^+) is shown on the left, and the reduced form (NADH) is shown on the right. The nitrogenous base in NADH has one more hydrogen ion and two more electrons than in NAD^+.

Electrons and Energy

The removal of an electron from a molecule (oxidizing it), results in a decrease in potential energy in the oxidized compound. The electron

(sometimes as part of a hydrogen atom) does not remain unbonded, however, in the cytoplasm of a cell. Rather, the electron is shifted to a second compound, reducing the second compound. *The shift of an electron from one compound to another removes some potential energy from the first compound (the oxidized compound) and increases the potential energy of the second compound (the reduced compound).* The transfer of electrons between molecules is important because most of the energy stored in atoms and used to fuel cell functions is in the form of high-energy electrons. The transfer of energy in the form of high-energy electrons allows the cell to transfer and use energy in an incremental fashion—in small packages rather than in a single, destructive burst. This chapter focuses on the extraction of energy from food; you will see that as you track the path of the transfers, you are tracking the path of electrons moving through metabolic pathways.

Electron Carriers

In living systems, a small class of compounds functions as electron shuttles: they bind and carry high-energy electrons between compounds in biochemical pathways. The principal electron carriers we will consider are derived from the B vitamin group and are derivatives of nucleotides. These compounds can be easily reduced (that is, they accept electrons) or oxidized (they lose electrons). Nicotinamide adenine dinucleotide

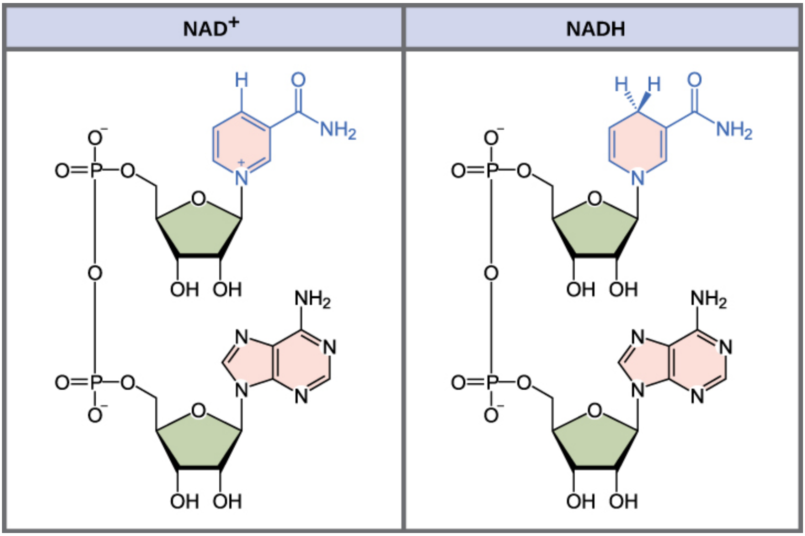
(NAD) ([link]) is derived from vitamin B_3, niacin. NAD^+ is the oxidized form of the molecule; NADH is the reduced form of the molecule after it has accepted two electrons and a proton (which together are the equivalent of a hydrogen atom with an extra electron). Note that if a compound has an "H" on it, it is generally reduced (e.g., NADH is the reduced form of NAD).

NAD^+ can accept electrons from an organic molecule according to the general equation:

$$\underset{\text{Reducing agent}}{\text{RH}} + \underset{\text{Oxidizing agent}}{\text{NAD}^+} \rightarrow \underset{\text{Reduced}}{\text{NADH}} + \underset{\text{Oxidized}}{\text{R}}$$

When electrons are added to a compound, *it is reduced*. A compound that reduces another is called a reducing agent. In the above equation, RH is a reducing agent, and NAD^+ is reduced to NADH. When electrons are removed from a compound, *it is oxidized*. A compound that oxidizes another is called an oxidizing agent. In the above equation, NAD^+ is an oxidizing agent, and RH is oxidized to R.

Similarly, flavin adenine dinucleotide (FAD^+) is derived from vitamin B_2, also called riboflavin. Its reduced form is $FADH_2$. A second variation of NAD, NADP, contains an extra phosphate group. Both NAD^+ and FAD^+ are extensively used in energy extraction from sugars, and NADP plays an important role in anabolic reactions and photosynthesis in plants.



ATP (adenosine triphosphate) has three phosphate groups that can be removed by hydrolysis (addition of H_2O) to form ADP (adenosine diphosphate) or AMP (adenosine monophosphate). The negative charges on the phosphate group naturally repel each other, requiring energy to bond them together and releasing energy when these bonds are broken. In phosphorylation reactions, the gamma (third) phosphate of ATP is attached to a protein. In eukaryotes, oxidative phosphorylation takes place in mitochondria. In prokaryotes, this process takes place in the plasma membrane. (Credit: modification of work by Mariana Ruiz Villareal)

ATP in Living Systems

A living cell cannot store significant amounts of free energy. Excess free energy would result in an increase of heat in the cell, which would result in

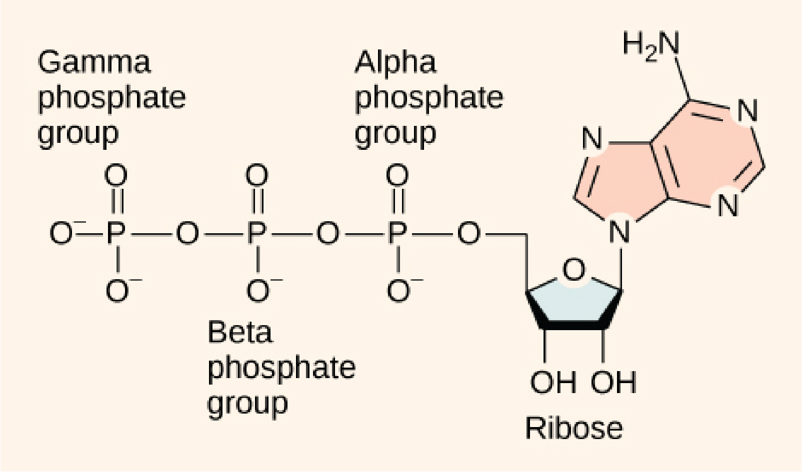
excessive thermal motion that could damage and then destroy the cell. Rather, a cell must be able to handle that energy in a way that enables the cell to store energy safely and release it for use only as needed. Living cells accomplish this by using the compound adenosine triphosphate (ATP). ATP is often called the “energy currency” of the cell, and, like currency, this versatile compound can be used to fill any energy need of the cell. How? It functions similarly to a rechargeable battery.

When ATP is broken down, usually by the removal of its terminal phosphate group, energy is released. The energy is used to do work by the cell, usually when the released phosphate binds to another molecule, thereby activating it. For example, in the mechanical work of muscle contraction, ATP supplies the energy to move the contractile muscle proteins. Recall the active transport work of the sodium-potassium pump in cell membranes. ATP alters the structure of the integral protein that functions as the pump, changing its affinity for sodium and potassium. In this way, the cell performs work, pumping ions against their electrochemical gradients.

ATP Structure and Function

At the heart of ATP is a molecule of adenosine monophosphate (AMP), which is composed of an adenine molecule bonded to a ribose molecule and

to a single phosphate group ([link]). Ribose is a five-carbon sugar found in RNA, and AMP is one of the nucleotides in RNA. The addition of a second phosphate group to this core molecule results in the formation of adenosine diphosphate (ADP); the addition of a third phosphate group forms adenosine triphosphate (ATP).



The addition of a phosphate group to a molecule requires energy. Phosphate groups are negatively charged and thus repel one another when they are arranged in series, as they are in ADP and ATP. This repulsion makes the ADP and ATP molecules inherently unstable. The release of one or two phosphate groups from ATP, a process called **dephosphorylation**, releases energy.

Energy from ATP

Hydrolysis is the process of breaking complex

macromolecules apart. During hydrolysis, water is split, or lysed, and the resulting hydrogen atom (H^+) and a hydroxyl group (OH^-), or *hydroxide,* are added to the larger molecule. The hydrolysis of ATP produces ADP, together with an inorganic phosphate ion (P_i), and the release of free energy. To carry out life processes, ATP is continuously broken down into ADP, and like a rechargeable battery, ADP is continuously regenerated into ATP by the reattachment of a third phosphate group. Water, which was broken down into its hydrogen atom and hydroxyl group (hydroxide) during ATP hydrolysis, is regenerated when a third phosphate is added to the ADP molecule, reforming ATP.

Obviously, energy must be infused into the system to regenerate ATP. Where does this energy come from? In nearly every living thing on Earth, the energy comes from the metabolism of glucose, fructose, or galactose, all isomers with the chemical formula $C_6H_{12}O_6$ but different molecular configurations. In this way, ATP is a direct link between the limited set of exergonic pathways of glucose catabolism and the multitude of endergonic pathways that power living cells.

Phosphorylation

Recall that, in some chemical reactions, enzymes may bind to several substrates that react with each other on the enzyme, forming an intermediate

complex. An intermediate complex is a temporary structure, and it allows one of the substrates (such as ATP) and reactants to more readily react with each other; in reactions involving ATP, ATP is one of the substrates and ADP is a product. During an endergonic chemical reaction, ATP forms an intermediate complex with the substrate and enzyme in the reaction. This intermediate complex allows the ATP to transfer its third phosphate group, with its energy, to the substrate, a process called phosphorylation. **Phosphorylation** refers to the addition of the phosphate (~P). This is illustrated by the following generic reaction, in which A and B represent two different substrates:

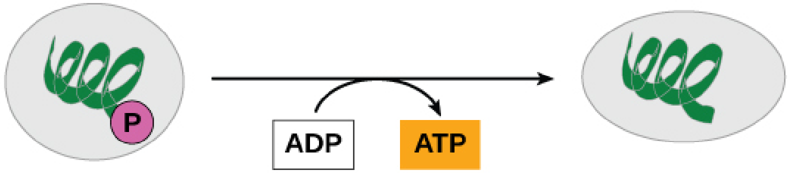
$$\text{A} + \text{enzyme} + \text{ATP} \rightarrow [\text{A} - \text{enzyme} - \sim\text{P}] \rightarrow \text{B} + \text{enzyme} + \text{ADP} + \text{phosphate ion}$$

When the intermediate complex breaks apart, the energy is used to modify the substrate and convert it into a product of the reaction. The ADP molecule and a free phosphate ion are released into the medium and are available for recycling through cell metabolism.

Substrate Phosphorylation

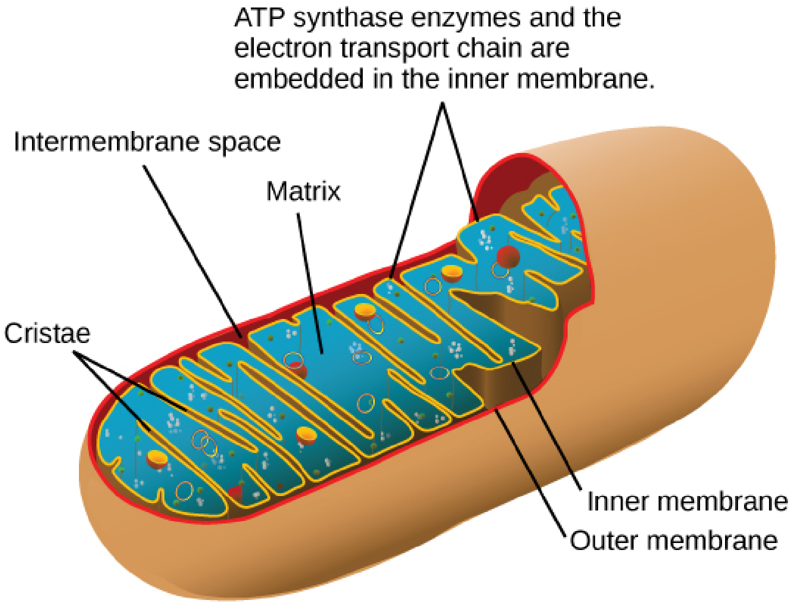
ATP is generated through two mechanisms during the breakdown of glucose. A few ATP molecules are generated (that is, regenerated from ADP) as a direct result of the chemical reactions that occur in the catabolic pathways. A phosphate group is

removed from an intermediate reactant in the pathway, and the free energy of the reaction is used to add the third phosphate to an available ADP molecule, producing ATP ([link]). This very direct method of phosphorylation is called **substrate-level phosphorylation**.



Oxidative Phosphorylation

Most of the ATP generated during glucose catabolism, however, is derived from a much more complex process, chemiosmosis, which takes place in mitochondria ([link]) within a eukaryotic cell or the plasma membrane of a prokaryotic cell. **Chemiosmosis**, a process of ATP production in cellular metabolism, is used to generate 90 percent of the ATP made during glucose catabolism and is also the method used in the light reactions of photosynthesis to harness the energy of sunlight. The production of ATP using the process of chemiosmosis is called **oxidative phosphorylation** because of the involvement of oxygen in the process.



Career Connections

Mitochondrial Disease Physician

What happens when the critical reactions of cellular respiration do not proceed correctly? This may happen in mitochondrial diseases, which are genetic disorders of metabolism. Mitochondrial disorders can arise from mutations in nuclear or mitochondrial DNA, and they result in the production of less energy than is normal in body cells. In type 2 diabetes, for instance, the oxidation efficiency of NADH is reduced, impacting oxidative phosphorylation but not the other steps of respiration. Symptoms of mitochondrial diseases can include muscle weakness, lack of coordination, stroke-like episodes, and loss of vision and hearing.

Most affected people are diagnosed in childhood, although there are some adult-onset diseases. Identifying and treating mitochondrial disorders is a specialized medical field. The educational preparation for this profession requires a college education, followed by medical school with a specialization in medical genetics. Medical geneticists can be board certified by the American Board of Medical Genetics and go on to become associated with professional organizations devoted to the study of mitochondrial diseases, such as the Mitochondrial Medicine Society and the Society for Inherited Metabolic Disorders.

Section Summary

ATP functions as the energy currency for cells. It allows the cell to store energy briefly and transport it within the cell to support endergonic chemical reactions. The structure of ATP is that of an RNA nucleotide with three phosphates attached. As ATP is used for energy, a phosphate group or two are detached, and either ADP or AMP is produced. Energy derived from glucose catabolism is used to convert ADP into ATP. When ATP is used in a reaction, the third phosphate is temporarily attached to a substrate in a process called

phosphorylation. The two processes of ATP regeneration that are used in conjunction with glucose catabolism are substrate-level phosphorylation and oxidative phosphorylation through the process of chemiosmosis.

Review Questions

The energy currency used by cells is _____.

1. ATP
2. ADP
3. AMP
4. adenosine

A

A reducing chemical reaction _____.

1. reduces the compound to a simpler form
2. adds an electron to the substrate
3. removes a hydrogen atom from the substrate
4. is a catabolic reaction

B

Critical Thinking Questions

Why is it beneficial for cells to use ATP rather than energy directly from the bonds of carbohydrates? What are the greatest drawbacks to harnessing energy directly from the bonds of several different compounds?

ATP provides the cell with a way to handle energy in an efficient manner. The molecule can be charged, stored, and used as needed. Moreover, the energy from hydrolyzing ATP is delivered as a consistent amount. Harvesting energy from the bonds of several different compounds would result in energy deliveries of different quantities.

Glossary

chemiosmosis
: process in which there is a production of adenosine triphosphate (ATP) in cellular metabolism by the involvement of a proton gradient across a membrane

dephosphorylation

removal of a phosphate group from a molecule

oxidative phosphorylation
: production of ATP using the process of chemiosmosis in the presence of oxygen

phosphorylation
: addition of a high-energy phosphate to a compound, usually a metabolic intermediate, a protein, or ADP

redox reaction
: chemical reaction that consists of the coupling of an oxidation reaction and a reduction reaction

substrate-level phosphorylation
: production of ATP from ADP using the excess energy from a chemical reaction and a phosphate group from a reactant

Regulation of Cellular Respiration

By the end of this section, you will be able to do the following:

- Describe how feedback inhibition would affect the production of an intermediate or product in a pathway
- Identify the mechanism that controls the rate of the transport of electrons through the electron transport chain

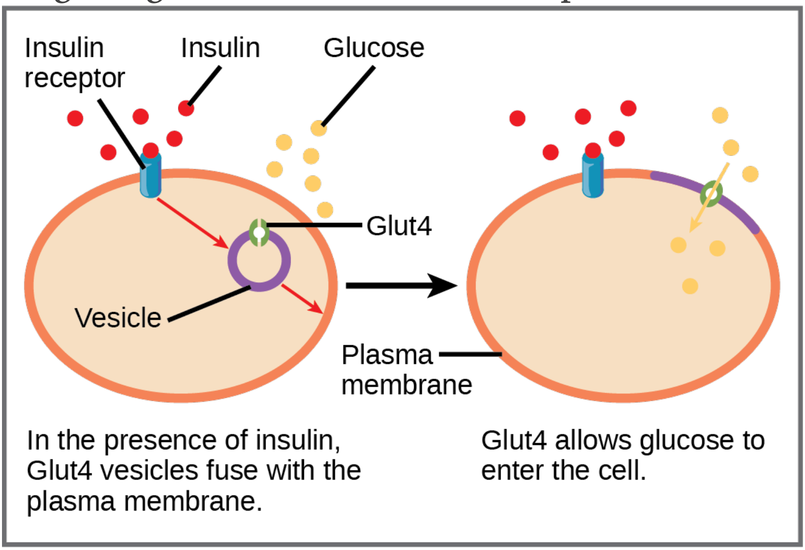
Cellular respiration must be regulated in order to provide balanced amounts of energy in the form of ATP. The cell also must generate a number of intermediate compounds that are used in the anabolism and catabolism of macromolecules. Without controls, metabolic reactions would quickly come to a standstill as the forward and backward reactions reached a state of equilibrium. Resources would be used inappropriately. A cell does not need the maximum amount of ATP that it can make all the time: At times, the cell needs to shunt some of the intermediates to pathways for amino acid, protein, glycogen, lipid, and nucleic acid production. In short, the cell needs to control its metabolism.

GLUT4 is a glucose transporter that is stored in vesicles. A cascade of events that occurs upon insulin binding to a receptor in the plasma membrane causes GLUT4-containing vesicles to fuse with the plasma membrane so that glucose may be

transported into the cell.

Regulatory Mechanisms

A variety of mechanisms is used to control cellular respiration. Some type of control exists at each stage of glucose metabolism. Access of glucose to the cell can be regulated using the **GLUT (glucose transporter) proteins** that transport glucose ([link]). Different forms of the GLUT protein control passage of glucose into the cells of specific tissues.



Some reactions are controlled by having two different enzymes—one each for the two directions of a reversible reaction. Reactions that are catalyzed by only one enzyme can go to equilibrium, stalling the reaction. In contrast, if two different enzymes (each specific for a given direction) are necessary for a reversible reaction, the opportunity to control

the rate of the reaction increases, and equilibrium is not reached.

A number of enzymes involved in each of the pathways—in particular, the enzyme catalyzing the first committed reaction of the pathway—are controlled by attachment of a molecule to an allosteric site on the protein. The molecules most commonly used in this capacity are the nucleotides ATP, ADP, AMP, NAD^+, and NADH. These regulators—allosteric effectors—may increase or decrease enzyme activity, depending on the prevailing conditions. The allosteric effector alters the steric structure of the enzyme, usually affecting the configuration of the active site. This alteration of the protein's (the enzyme's) structure either increases or decreases its affinity for its substrate, with the effect of increasing or decreasing the rate of the reaction. The attachment signals to the enzyme. This binding can increase or decrease the enzyme's activity, providing a feedback mechanism. This feedback type of control is effective as long as the chemical affecting it is attached to the enzyme. Once the overall concentration of the chemical decreases, it will diffuse away from the protein, and the control is relaxed.

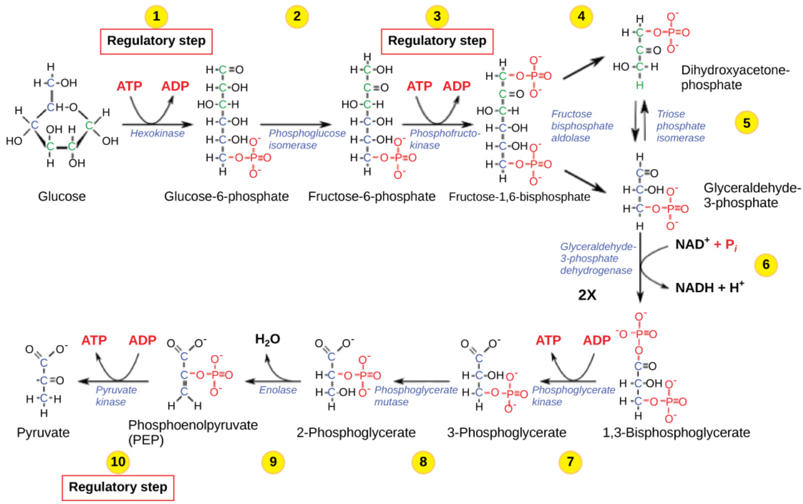
The glycolysis pathway is primarily regulated at the three key enzymatic steps (1, 3, and 10) as indicated. Note that the first two steps that are regulated occur early in the pathway and involve hydrolysis of ATP.

Control of Catabolic Pathways

Enzymes, proteins, electron carriers, and pumps that play roles in glycolysis, the citric acid cycle, and the electron transport chain tend to catalyze nonreversible reactions. In other words, if the initial reaction takes place, the pathway is committed to proceeding with the remaining reactions. Whether a particular enzyme activity is released depends upon the energy needs of the cell (as reflected by the levels of ATP, ADP, and AMP).

Glycolysis

The control of glycolysis begins with the first enzyme in the pathway, hexokinase ([link]). This enzyme catalyzes the phosphorylation of glucose, which helps to prepare the compound for cleavage in a later step. The presence of the negatively charged phosphate in the molecule also prevents the sugar from leaving the cell. When hexokinase is inhibited, glucose diffuses out of the cell and does not become a substrate for the respiration pathways in that tissue. The product of the hexokinase reaction is glucose-6-phosphate, which accumulates when a later enzyme, phosphofructokinase, is inhibited.



Phosphofructokinase is the main enzyme controlled in glycolysis. High levels of ATP or citrate or a lower, more acidic pH decreases the enzyme's activity. An increase in citrate concentration can occur because of a blockage in the citric acid cycle. Fermentation, with its production of organic acids such as lactic acid, frequently accounts for the increased acidity in a cell; however, the products of fermentation do not typically accumulate in cells.

The last step in glycolysis is catalyzed by pyruvate kinase. The pyruvate produced can proceed to be catabolized or converted into the amino acid alanine. If no more energy is needed and alanine is in adequate supply, the enzyme is inhibited. The enzyme's activity is increased when fructose-1,6-bisphosphate levels increase. (Recall that fructose-1,6-bisphosphate is an intermediate in the first half of glycolysis.) The regulation of pyruvate

kinase involves phosphorylation by a kinase (pyruvate kinase), resulting in a less-active enzyme. Dephosphorylation by a phosphatase reactivates it. Pyruvate kinase is also regulated by ATP (a negative allosteric effect).

If more energy is needed, more pyruvate will be converted into acetyl CoA through the action of pyruvate dehydrogenase. If either acetyl groups or NADH accumulates, there is less need for the reaction, and the rate decreases. Pyruvate dehydrogenase is also regulated by phosphorylation: a kinase phosphorylates it to form an inactive enzyme, and a phosphatase reactivates it. The kinase and the phosphatase are also regulated.

Citric Acid Cycle

The citric acid cycle is controlled through the enzymes that catalyze the reactions that make the first two molecules of NADH ([link]). These enzymes are isocitrate dehydrogenase and *α*-ketoglutarate dehydrogenase. When adequate ATP and NADH levels are available, the rates of these reactions decrease. When more ATP is needed, as reflected in rising ADP levels, the rate increases. Alpha-ketoglutarate dehydrogenase will also be affected by the levels of succinyl CoA—a subsequent intermediate in the cycle—causing a decrease in activity. A decrease in the rate of operation of the pathway at this point is not necessarily negative, as

the increased levels of the α-ketoglutarate not used by the citric acid cycle can be used by the cell for amino acid (glutamate) synthesis.

Electron Transport Chain

Specific enzymes of the electron transport chain are unaffected by feedback inhibition, but the rate of electron transport through the pathway is affected by the levels of ADP and ATP. Greater ATP consumption by a cell is indicated by a buildup of ADP. As ATP usage decreases, the concentration of ADP decreases, and now, ATP begins to build up in the cell. This change in the relative concentration of ADP to ATP triggers the cell to slow down the electron transport chain.

Link to Learning

Visit this site to see an animation of the electron transport chain and ATP synthesis.

For a summary of feedback controls in cellular respiration, see [link].

Summary of Feedback Controls in Cellular Respiration			
Pathway	Enzyme affected	Elevated levels of effector	Effect on pathway activity
glycolysis	hexokinase	glucose-6-phosphate	decrease
	phosphofructokinase	low-energy charge (ATP, AMP), fructose-6-phosphate via fructose-2,6-bisphosphate	increase
		high-energy charge (ATP, AMP), citrate, acidic pH	decrease
	pyruvate kinase	fructose-1,6-bisphosphate	increase
		high-energy charge (ATP, AMP), alanine	decrease
pyruvate to acetyl CoA	pyruvate dehydrogenase	ADP, pyruvate	increase

conversion			
		acetyl CoA, ATP, NADH	decrease
citric acid cycle	isocitrate dehydrogenase	ADP	increase
		ATP, NADH	decrease
	α-ketoglutarate dehydrogenase	calcium ions, ADP	increase
		ATP, NADH, succinyl CoA	decrease
electron transport chain		ADP	increase
		ATP	decrease

Section Summary

Cellular respiration is controlled by a variety of means. The entry of glucose into a cell is controlled by the transport proteins that aid glucose passage through the cell membrane. Most of the control of the respiration processes is accomplished through the control of specific enzymes in the pathways. This is a type of negative feedback mechanism, turning the enzymes off. The enzymes respond most often to the levels of the available nucleosides ATP, ADP, AMP, NAD^+, and FAD. Other intermediates of the pathway also affect certain enzymes in the

systems.

Review Questions

The effect of high levels of ADP is to ______ in cellular respiration.

1. increase the activity of specific enzymes
2. decrease the activity of specific enzymes
3. have no effect on the activity of specific enzymes
4. slow down the pathway

A

The control of which enzyme exerts the most control on glycolysis?

1. hexokinase
2. phosphofructokinase
3. glucose-6-phosphatase
4. aldolase

B

Critical Thinking Questions

How does citrate from the citric acid cycle affect glycolysis?

Citrate can inhibit phosphofructokinase by feedback regulation.

Why might negative feedback mechanisms be more common than positive feedback mechanisms in living cells?

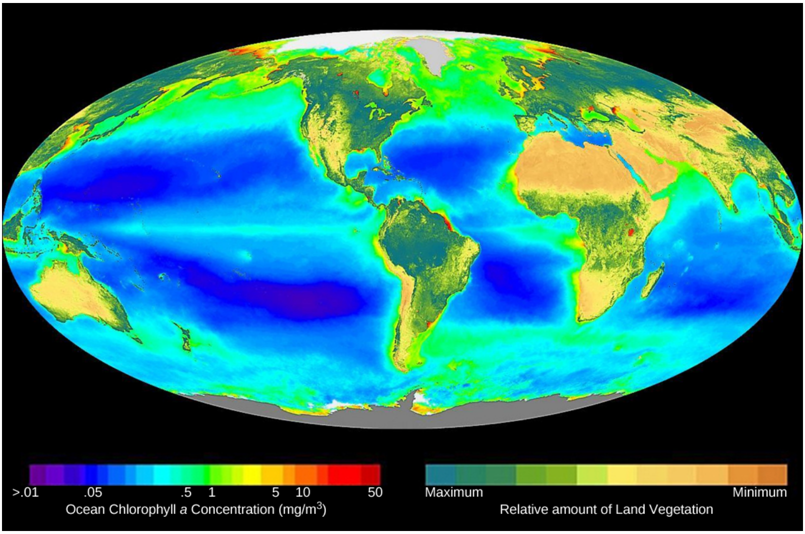
Negative feedback mechanisms actually control a process; it can turn it off, whereas positive feedback accelerates the process, allowing the cell no control over it. Negative feedback naturally maintains homeostasis, whereas positive feedback drives the system away from equilibrium.

Glossary

GLUT protein
: integral membrane protein that transports glucose

Introduction

class="introduction" This world map shows Earth’s distribution of photosynthetic activity determined by chlorophyll *a* concentrations. On land, chlorophyll is evident from terrestrial plants, and within oceanic zones, from chlorophyll from phytoplankton. (credit: modification of work by SeaWiFS Project, NASA/Goddard Space Flight Center and ORBIMAGE)



The metabolic processes in all organisms—from bacteria to humans—require energy. To get this energy, many organisms access stored energy by eating, that is, by ingesting other organisms. But where does the stored energy in food originate? All of this energy can be traced back to photosynthesis.

Overview of Photosynthesis

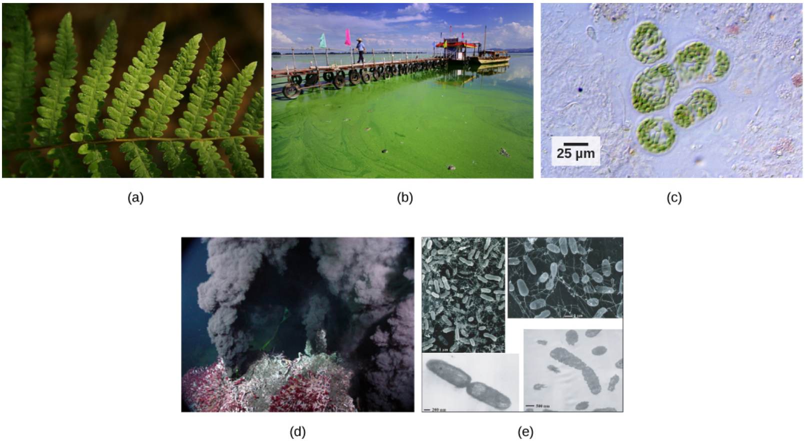
By the end of this section, you will be able to do the following:

- Explain the significance of photosynthesis to other living organisms
- Describe the main structures involved in photosynthesis
- Identify the substrates and products of photosynthesis

Photosynthesis is essential to all life on earth; both plants and animals depend on it. It is the only biological process that can capture energy that originates from sunlight and converts it into chemical compounds (carbohydrates) that every organism uses to power its metabolism. It is also a source of oxygen necessary for many living organisms. In brief, the energy of sunlight is "captured" to energize electrons, whose energy is then stored in the covalent bonds of sugar molecules. How long lasting and stable are those covalent bonds? The energy extracted today by the burning of coal and petroleum products represents sunlight energy captured and stored by photosynthesis 350 to 200 million years ago during the Carboniferous Period.

Plants, algae, and a group of bacteria called cyanobacteria are the only organisms capable of performing photosynthesis ([link]). Because they

use light to manufacture their own food, they are called **photoautotrophs** (literally, “self-feeders using light”). Other organisms, such as animals, fungi, and most other bacteria, are termed **heterotrophs** (“other feeders”), because they must rely on the sugars produced by photosynthetic organisms for their energy needs. A third very interesting group of bacteria synthesize sugars, not by using sunlight’s energy, but by extracting energy from inorganic chemical compounds. For this reason, they are referred to as **chemoautotrophs**. Photoautotrophs including (a) plants, (b) algae, and (c) cyanobacteria synthesize their organic compounds via photosynthesis using sunlight as an energy source. Cyanobacteria and planktonic algae can grow over enormous areas in water, at times completely covering the surface. In a (d) deep sea vent, chemoautotrophs, such as these (e) thermophilic bacteria, capture energy from inorganic compounds to produce organic compounds. The ecosystem surrounding the vents has a diverse array of animals, such as tubeworms, crustaceans, and octopuses that derive energy from the bacteria. (credit a: modification of work by Steve Hillebrand, U.S. Fish and Wildlife Service; credit b: modification of work by "eutrophication&hypoxia"/Flickr; credit c: modification of work by NASA; credit d: University of Washington, NOAA; credit e: modification of work by Mark Amend, West Coast and Polar Regions Undersea Research Center, UAF, NOAA)



The importance of photosynthesis is not just that it can capture sunlight's energy. After all, a lizard sunning itself on a cold day can use the sun's energy to warm up in a process called *behavioral thermoregulation*. In contrast, photosynthesis is vital because it evolved as a way to *store the energy from solar radiation (the "photo-" part) to energy in the carbon-carbon bonds of carbohydrate molecules (the "-synthesis" part)*. Those carbohydrates are the energy source that heterotrophs use to power the synthesis of ATP via respiration. Therefore, photosynthesis powers 99 percent of Earth's ecosystems. When a top predator, such as a wolf, preys on a deer ([link]), the wolf is at the end of an energy path that went from nuclear reactions on the surface of the sun, to visible light, to photosynthesis, to vegetation, to deer, and finally to the wolf.

The energy stored in carbohydrate molecules from photosynthesis passes through the food chain. The predator that eats these deer receives a portion of the energy that originated in the photosynthetic

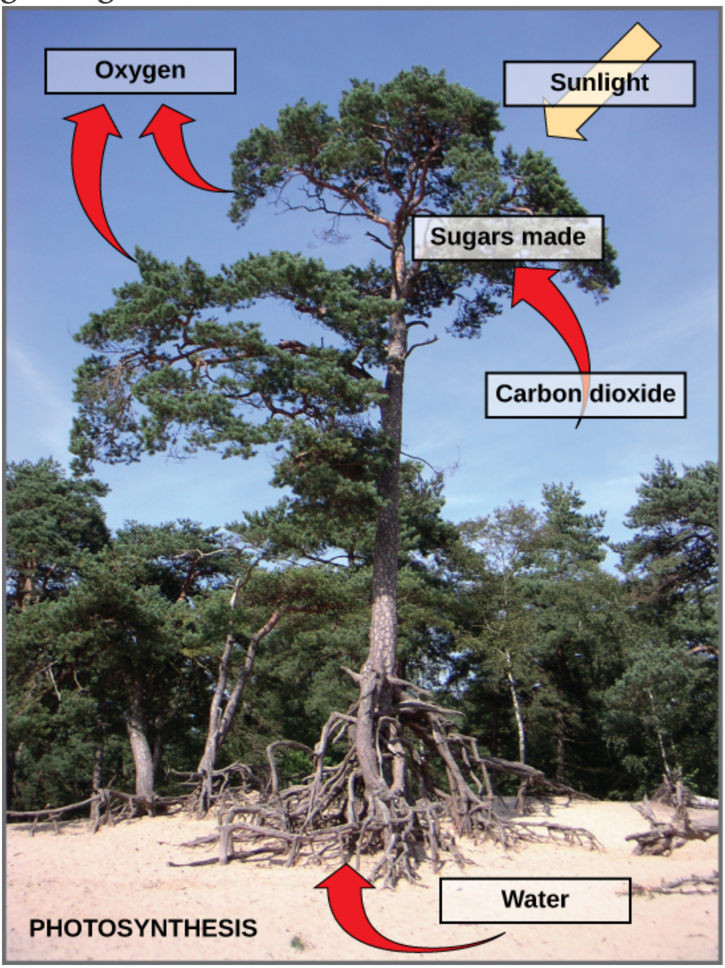
vegetation that the deer consumed. (credit: modification of work by Steve VanRiper, U.S. Fish and Wildlife Service)

Photosynthesis uses solar energy, carbon dioxide, and water to produce energy-storing carbohydrates. Oxygen is generated as a waste product of photosynthesis. The basic equation for photosynthesis is deceptively simple. In reality, the process takes place in many steps involving intermediate reactants and products. Glucose, the primary energy source in cells, is made from two three-carbon GA3Ps.

Main Structures and Summary of Photosynthesis

Photosynthesis is a multi-step process that requires specific wavelengths of visible sunlight, carbon dioxide (which is low in energy), and water as

substrates ([link]). After the process is complete, it releases oxygen and produces glyceraldehyde-3-phosphate (GA3P), as well as simple carbohydrate molecules (high in energy) that can then be converted into glucose, sucrose, or any of dozens of other sugar molecules. These sugar molecules contain energy and the energized carbon that all living things need to survive.



The following is the chemical equation for

photosynthesis ([link]):

Photosynthesis Equation			
Carbon dioxide +	Water $\xrightarrow{\text{SUNLIGHT}}$	Sugar +	Oxygen
$6CO_2$	$6H_2O$	$C_6H_{12}O_6$	$6O_2$

Although the equation looks simple, the many steps that take place during photosynthesis are actually quite complex. Before learning the details of how photoautotrophs turn sunlight into food, it is important to become familiar with the structures involved.

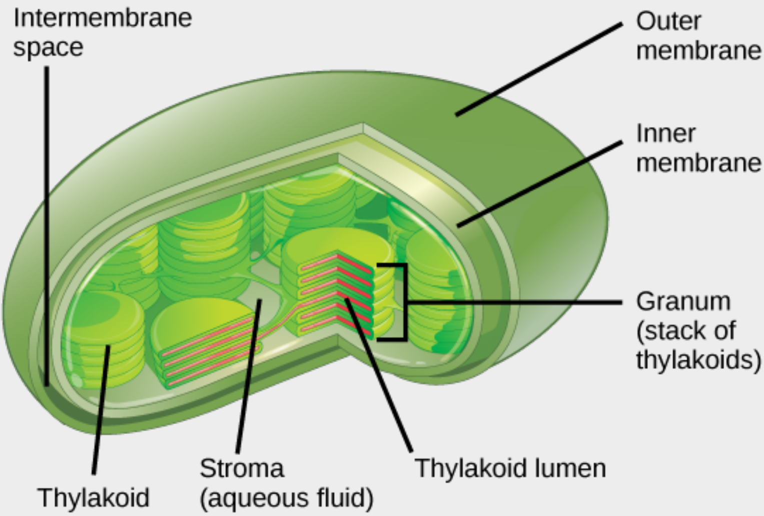
Basic Photosynthetic Structures

In plants, photosynthesis generally takes place in leaves, which consist of several layers of cells. The process of photosynthesis occurs in a middle layer called the **mesophyll**. The gas exchange of carbon dioxide and oxygen occurs through small, regulated openings called **stomata** (singular: stoma), which also play roles in the regulation of gas exchange and water balance. The stomata are typically located on the underside of the leaf, which helps to minimize water loss due to high temperatures on the upper surface of the leaf. Each stoma is flanked by guard cells that regulate the opening and closing of the stomata by swelling or shrinking in response to osmotic changes.

In all autotrophic eukaryotes, photosynthesis takes place inside an organelle called a **chloroplast**. For plants, chloroplast-containing cells exist mostly in the mesophyll. Chloroplasts have a double membrane envelope (composed of an outer membrane and an inner membrane), and are ancestrally derived from ancient free-living cyanobacteria. Within the chloroplast are stacked, disc-shaped structures called **thylakoids**. Embedded in the thylakoid membrane is chlorophyll, a **pigment** (molecule that absorbs light) responsible for the initial interaction between light and plant material, and numerous proteins that make up the electron transport chain. The thylakoid membrane encloses an internal space called the **thylakoid lumen**. As shown in [link], a stack of thylakoids is called a **granum**, and the liquid-filled space surrounding the granum is called **stroma** or "bed" (not to be confused with stoma or "mouth," an opening on the leaf epidermis).

Visual Connection

Photosynthesis takes place in chloroplasts, which have an outer membrane and an inner membrane. Stacks of thylakoids called grana form a third membrane layer.



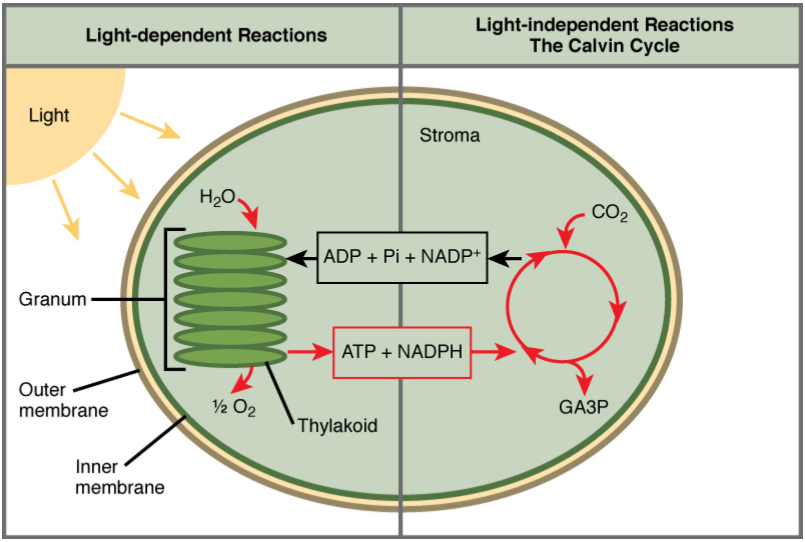
On a hot, dry day, the guard cells of plants close their stomata to conserve water. What impact will this have on photosynthesis?

Photosynthesis takes place in two stages: light-dependent reactions and the Calvin cycle. Light-dependent reactions, which take place in the thylakoid membrane, use light energy to make ATP and NADPH. The Calvin cycle, which takes place in the stroma, uses energy derived from these compounds to make GA3P from CO_2.

The Two Parts of Photosynthesis

Photosynthesis takes place in two sequential stages: the light-dependent reactions and the light-independent reactions. In the **light-dependent**

reactions, energy from sunlight is absorbed by chlorophyll and that energy is converted into stored chemical energy. In the **light-independent reactions**, the chemical energy harvested during the light-dependent reactions drives the assembly of sugar molecules from carbon dioxide. Therefore, although the light-independent reactions do not use light as a reactant, they require the products of the light-dependent reactions to function. In addition, however, several enzymes of the light-independent reactions are activated by light. The light-dependent reactions utilize certain molecules to temporarily store the energy: These are referred to as *energy carriers*. The energy carriers that move energy from light-dependent reactions to light-independent reactions can be thought of as “full” because they are rich in energy. After the energy is released, the “empty” energy carriers return to the light-dependent reaction to obtain more energy. [link] illustrates the components inside the chloroplast where the light-dependent and light-independent reactions take place.



Link to Learning

Click the link to learn more about photosynthesis.

Everyday Connection

Photosynthesis at the Grocery Store

Foods that humans consume originate from photosynthesis. (credit: Associação Brasileira de Supermercados)

Major grocery stores in the United States are organized into departments, such as dairy, meats, produce, bread, cereals, and so forth. Each aisle ([link]) contains hundreds, if not thousands, of different products for customers to buy and consume.

Although there is a large variety, each item ultimately can be linked back to photosynthesis. Meats and dairy link, because the animals were fed plant-based foods. The breads, cereals, and pastas come largely from starchy grains, which are the seeds of photosynthesis-dependent plants. What about desserts and drinks? All of these products contain sugar—sucrose is a plant product, a disaccharide, a carbohydrate molecule, which is built directly from photosynthesis. Moreover, many items are less obviously derived from plants: For instance, paper goods are generally plant products, and many plastics (abundant as products and packaging) are derived from “algae” (unicellular plant-like organisms, and cyanobacteria). Virtually

every spice and flavoring in the spice aisle was produced by a plant as a leaf, root, bark, flower, fruit, or stem. Ultimately, photosynthesis connects to every meal and every food a person consumes.

Section Summary

The process of photosynthesis transformed life on Earth. By harnessing energy from the sun, the evolution of photosynthesis allowed living things access to enormous amounts of energy. Because of photosynthesis, living things gained access to sufficient energy that allowed them to build new structures and achieve the biodiversity evident today.

Only certain organisms (photoautotrophs), can perform photosynthesis; they require the presence of chlorophyll, a specialized pigment that absorbs certain wavelengths of the visible spectrum and can capture energy from sunlight. Photosynthesis uses carbon dioxide and water to assemble carbohydrate molecules and release oxygen as a byproduct into the atmosphere. Eukaryotic autotrophs, such as plants and algae, have organelles called chloroplasts in which photosynthesis takes place, and starch accumulates. In prokaryotes, such as cyanobacteria,

the process is less localized and occurs within folded membranes, extensions of the plasma membrane, and in the cytoplasm.

Visual Connection Questions

[link] On a hot, dry day, the guard cells of plants close their stomata to conserve water. What impact will this have on photosynthesis?

[link] Levels of carbon dioxide (a necessary photosynthetic substrate) will immediately fall. As a result, the rate of photosynthesis will be inhibited.

Review Questions

Which of the following components is *not* used by both plants and cyanobacteria to carry out photosynthesis?

1. chloroplasts
2. chlorophyll
3. carbon dioxide

4. water

A

What two main products result from photosynthesis?

1. oxygen and carbon dioxide
2. chlorophyll and oxygen
3. sugars/carbohydrates and oxygen
4. sugars/carbohydrates and carbon dioxide

C

In which compartment of the plant cell do the light-independent reactions of photosynthesis take place?

1. thylakoid
2. stroma
3. outer membrane
4. mesophyll

B

Which statement about thylakoids in eukaryotes

is *not* correct?

1. Thylakoids are assembled into stacks.
2. Thylakoids exist as a maze of folded membranes.
3. The space surrounding thylakoids is called stroma.
4. Thylakoids contain chlorophyll.

B

Predict the end result if a chloroplast’s light-independent enzymes developed a mutation that prevented them from activating in response to light.

1. GA3P accumulation
2. ATP and NADPH accumulation
3. Water accumulation
4. Carbon dioxide depletion

B

How are the NADPH and GA3P molecules made during photosynthesis similar?

1. They are both end products of photosynthesis.

2. They are both substrates for photosynthesis.
3. They are both produced from carbon dioxide.
4. They both store energy in chemical bonds.

D

Critical Thinking Questions

What is the overall outcome of the light reactions in photosynthesis?

The outcome of light reactions in photosynthesis is the conversion of solar energy into chemical energy that the chloroplasts can use to do work (mostly anabolic production of carbohydrates from carbon dioxide).

Why are carnivores, such as lions, dependent on photosynthesis to survive?

Because lions eat animals that eat plants.

Why are energy carriers thought of as either “full” or “empty”?

The energy carriers that move from the light-dependent reaction to the light-independent one are “full” because they bring energy. After the energy is released, the “empty” energy carriers return to the light-dependent reaction to obtain more energy. There is not much actual movement involved. Both ATP and NADPH are produced in the stroma where they are also used and reconverted into ADP, Pi, and NADP+.

Describe how the grey wolf population would be impacted by a volcanic eruption that spewed a dense ash cloud that blocked sunlight in a section of Yellowstone National Park.

The grey wolves are apex predators in their food web, meaning they consume smaller prey animals and are not the prey of any other animal. Blocking sunlight would prevent the plants at the bottom of the food web from performing photosynthesis. This would kill many of the plants, reducing the food sources available to smaller animals in Yellowstone. A smaller prey animal population means that fewer wolves can survive in the area, and the

population of grey wolves will decrease.

How does the closing of the stomata limit photosynthesis?

The stomata regulate the exchange of gases and water vapor between a leaf and its surrounding environment. When the stomata are closed, the water molecules cannot escape the leaf, but the leaf also cannot acquire new carbon dioxide molecules from the environment. This limits the light-independent reactions to only continuing until the carbon dioxide stores in the leaf are depleted.

Glossary

chemoautotroph
: organism that can build organic molecules using energy derived from inorganic chemicals instead of sunlight

chloroplast
: organelle in which photosynthesis takes place

granum
: stack of thylakoids located inside a chloroplast

heterotroph
: organism that consumes organic substances or other organisms for food

light-dependent reaction
: first stage of photosynthesis where certain wavelengths of the visible light are absorbed to form two energy-carrying molecules (ATP and NADPH)

light-independent reaction
: second stage of photosynthesis, through which carbon dioxide is used to build carbohydrate molecules using energy from ATP and NADPH

mesophyll
: middle layer of chlorophyll-rich cells in a leaf

photoautotroph
: organism capable of producing its own organic compounds from sunlight

pigment
: molecule that is capable of absorbing certain wavelengths of light and reflecting others (which accounts for its color)

stoma
: opening that regulates gas exchange and water evaporation between leaves and the environment, typically situated on the

underside of leaves

stroma

fluid-filled space surrounding the grana inside a chloroplast where the light-independent reactions of photosynthesis take place

thylakoid

disc-shaped, membrane-bound structure inside a chloroplast where the light-dependent reactions of photosynthesis take place; stacks of thylakoids are called grana

thylakoid lumen

aqueous space bound by a thylakoid membrane where protons accumulate during light-driven electron transport

The Light-Dependent Reactions of Photosynthesis

By the end of this section, you will be able to do the following:

- Explain how plants absorb energy from sunlight
- Describe short and long wavelengths of light
- Describe how and where photosynthesis takes place within a plant

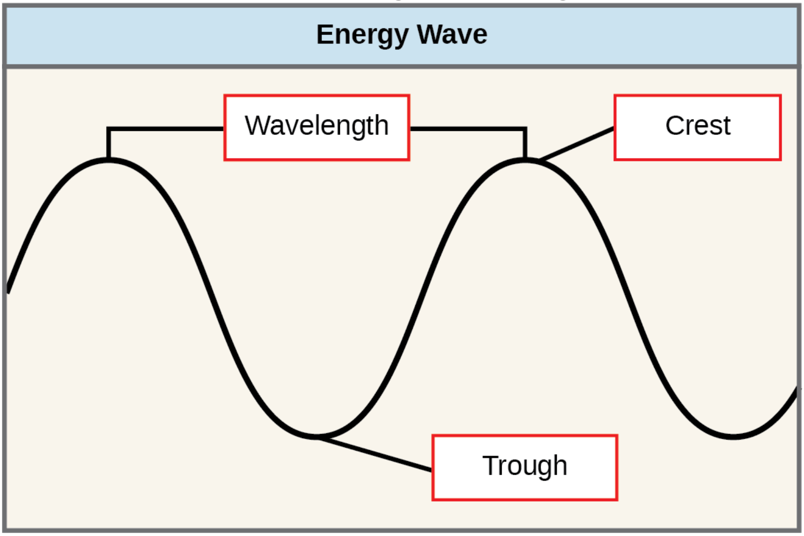
How can light energy be used to make food? When a person turns on a lamp, electrical energy becomes light energy. Like all other forms of kinetic energy, light can travel, change form, and be harnessed to do work. In the case of photosynthesis, light energy is converted into chemical energy, which photoautotrophs use to build basic carbohydrate molecules ([link]). However, autotrophs only use a few specific wavelengths of sunlight.

Photoautotrophs can capture visible light energy in specific wavelengths from the sun, converting it into the chemical energy used to build food molecules. (credit: Gerry Atwell)

The wavelength of a single wave is the distance between two consecutive points of similar position (two crests or two troughs) along the wave. The sun emits energy in the form of electromagnetic radiation. This radiation exists at different wavelengths, each of which has its own characteristic energy. All electromagnetic radiation, including visible light, is characterized by its wavelength.

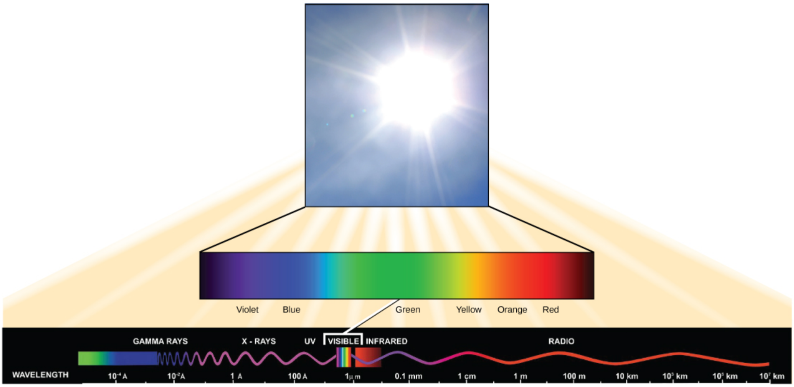
What Is Light Energy?

The sun emits an enormous amount of electromagnetic radiation (solar energy in a spectrum from very short gamma rays to very long radio waves). Humans can see only a tiny fraction of this energy, which we refer to as “visible light.” The manner in which solar energy travels is described as waves. Scientists can determine the amount of energy of a wave by measuring its **wavelength** (shorter wavelengths are more powerful than longer wavelengths)—the distance between consecutive crest points of a wave. Therefore, a single wave is measured from two consecutive points, such as from crest to crest or from trough to trough ([link]).



Visible light constitutes only one of many types of electromagnetic radiation emitted from the sun and other stars. Scientists differentiate the various types

of radiant energy from the sun within the electromagnetic spectrum. The **electromagnetic spectrum** is the range of all possible frequencies of radiation ([link]). The difference between wavelengths relates to the amount of energy carried by them.



Each type of electromagnetic radiation travels at a particular wavelength. The longer the wavelength, the less energy it carries. Short, tight waves carry the most energy. This may seem illogical, but think of it in terms of a piece of moving heavy rope. It takes little effort by a person to move a rope in long, wide waves. To make a rope move in short, tight waves, a person would need to apply significantly more energy.

The electromagnetic spectrum ([link]) shows several types of electromagnetic radiation originating from the sun, including X-rays and ultraviolet (UV) rays. The higher-energy waves can penetrate tissues and damage cells and DNA, which explains why both X-rays and UV rays can be harmful to living

organisms.

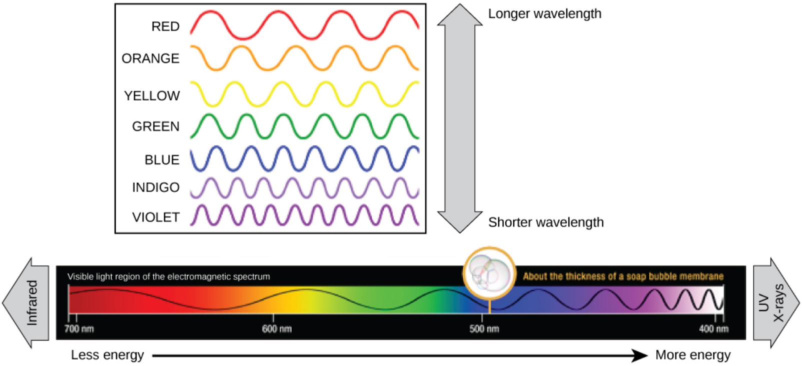
The colors of visible light do not carry the same amount of energy. Violet has the shortest wavelength and therefore carries the most energy, whereas red has the longest wavelength and carries the least amount of energy. (credit: modification of work by NASA) (a) Chlorophyll *a*, (b) chlorophyll *b*, and (c) *β*-carotene are hydrophobic organic pigments found in the thylakoid membrane. Chlorophyll *a* and *b*, which are identical except for the part indicated in the red box, are responsible for the green color of leaves. *β*-carotene is responsible for the orange color in carrots. Each pigment has (d) a unique absorbance spectrum. Plants that commonly grow in the shade have adapted to low levels of light by changing the relative concentrations of their chlorophyll pigments. (credit: Jason Hollinger)

Absorption of Light

Light energy initiates the process of photosynthesis when pigments absorb specific wavelengths of visible light. Organic pigments, whether in the human retina or the chloroplast thylakoid, have a narrow range of energy levels that they can absorb. Energy levels lower than those represented by red light are insufficient to raise an orbital electron to a excited (quantum) state. Energy levels higher than those in blue light will physically tear the molecules apart, in a process called bleaching. Our retinal

pigments can only "see" (absorb) wavelengths between 700 nm and 400 nm of light, a spectrum that is therefore called visible light. For the same reasons, plants, pigment molecules absorb only light in the wavelength range of 700 nm to 400 nm; plant physiologists refer to this range for plants as photosynthetically active radiation.

The visible light seen by humans as white light actually exists in a rainbow of colors. Certain objects, such as a prism or a drop of water, disperse white light to reveal the colors to the human eye. The visible light portion of the electromagnetic spectrum shows the rainbow of colors, with violet and blue having shorter wavelengths, and therefore higher energy. At the other end of the spectrum toward red, the wavelengths are longer and have lower energy ([link]).



Understanding Pigments

Different kinds of pigments exist, and each absorbs only specific wavelengths (colors) of visible light.

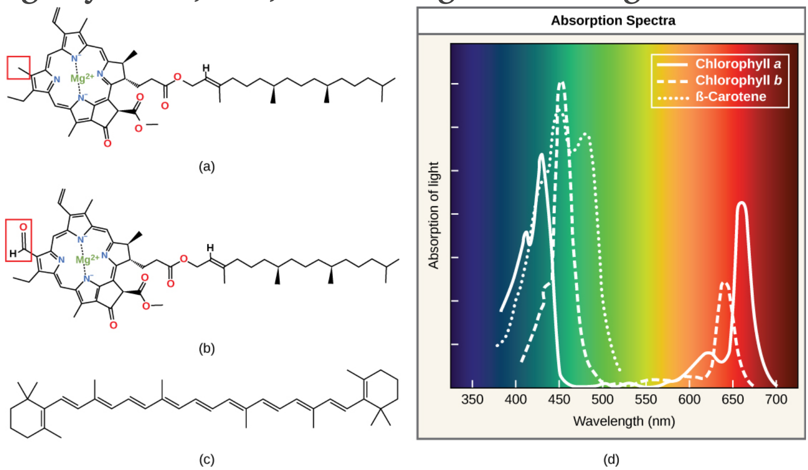
Pigments reflect or transmit the wavelengths they cannot absorb, making them appear a mixture of the reflected or transmitted light colors.

Chlorophylls and carotenoids are the two major classes of photosynthetic pigments found in plants and algae; each class has multiple types of pigment molecules. There are five major chlorophylls: *a*, *b*, *c* and *d* and a related molecule found in prokaryotes called *bacteriochlorophyll*. **Chlorophyll *a*** and **chlorophyll *b*** are found in higher plant chloroplasts and will be the focus of the following discussion.

With dozens of different forms, carotenoids are a much larger group of pigments. The carotenoids found in fruit—such as the red of tomato (lycopene), the yellow of corn seeds (zeaxanthin), or the orange of an orange peel (β-carotene)—are used as advertisements to attract seed dispersers. In photosynthesis, **carotenoids** function as photosynthetic pigments that are very efficient molecules for the disposal of excess energy. When a leaf is exposed to full sun, the light-dependent reactions are required to process an enormous amount of energy; if that energy is not handled properly, it can do significant damage. Therefore, many carotenoids reside in the thylakoid membrane, absorb excess energy, and safely dissipate that energy as heat.

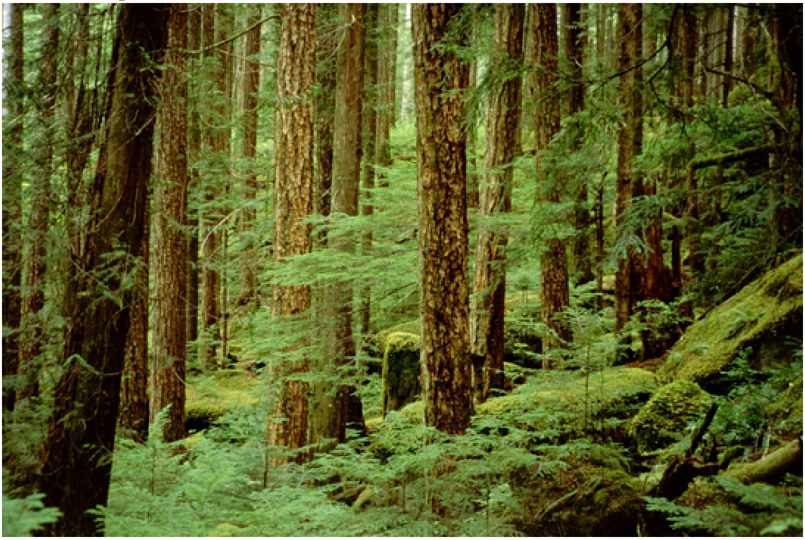
Each type of pigment can be identified by the

specific pattern of wavelengths it absorbs from visible light: This is termed the **absorption spectrum**. The graph in [link] shows the absorption spectra for chlorophyll *a*, chlorophyll *b*, and a type of carotenoid pigment called β-carotene (which absorbs blue and green light). Notice how each pigment has a distinct set of peaks and troughs, revealing a highly specific pattern of absorption. Chlorophyll *a* absorbs wavelengths from either end of the visible spectrum (blue and red), but not green. Because green is reflected or transmitted, chlorophyll appears green. Carotenoids absorb in the short-wavelength blue region, and reflect the longer yellow, red, and orange wavelengths.



Many photosynthetic organisms have a mixture of pigments, and by using these pigments, the organism can absorb energy from a wider range of wavelengths. Not all photosynthetic organisms have full access to sunlight. Some organisms grow underwater where light intensity and quality

decrease and change with depth. Other organisms grow in competition for light. Plants on the rainforest floor must be able to absorb any bit of light that comes through, because the taller trees absorb most of the sunlight and scatter the remaining solar radiation ([link]).



When studying a photosynthetic organism, scientists can determine the types of pigments present by generating absorption spectra. An instrument called a **spectrophotometer** can differentiate which wavelengths of light a substance can absorb. Spectrophotometers measure transmitted light and compute from it the absorption. By extracting pigments from leaves and placing these samples into a spectrophotometer, scientists can identify which wavelengths of light an organism can absorb. Additional methods for the identification of plant pigments include various types of chromatography

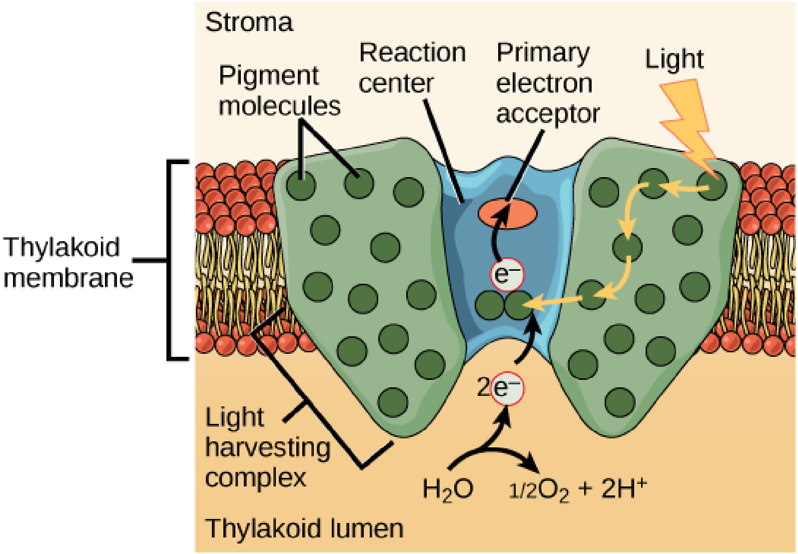
that separate the pigments by their relative affinities to solid and mobile phases.

A photosystem consists of 1) a light-harvesting complex and 2) a reaction center. Pigments in the light-harvesting complex pass light energy to two special chlorophyll *a* molecules in the reaction center. The light excites an electron from the chlorophyll *a* pair, which passes to the primary electron acceptor. The excited electron must then be replaced. In (a) photosystem II, the electron comes from the splitting of water, which releases oxygen as a waste product. In (b) photosystem I, the electron comes from the chloroplast electron transport chain discussed below.

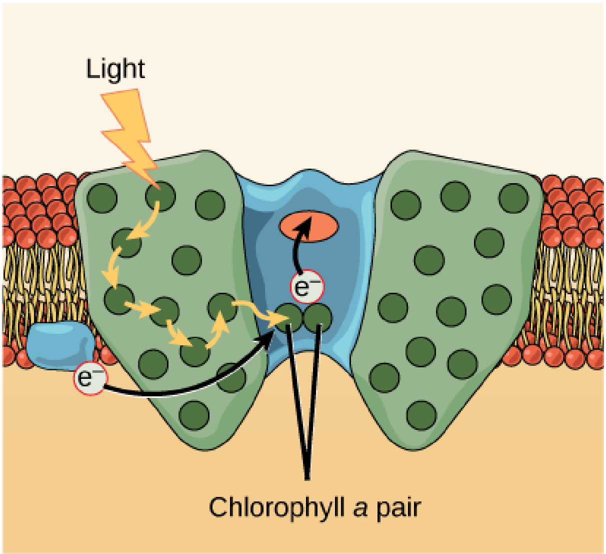
How Light-Dependent Reactions Work

The overall function of light-dependent reactions is to convert solar energy into chemical energy in the form of NADPH and ATP. This chemical energy supports the light-independent reactions and fuels the assembly of sugar molecules. The light-dependent reactions are depicted in [link]. Protein complexes and pigment molecules work together to produce NADPH and ATP. The numbering of the photosystems is derived from the order in which they were discovered, not in the order of the transfer of electrons.

(a) Photosystem II (P680)



(b) Photosystem I (P700)



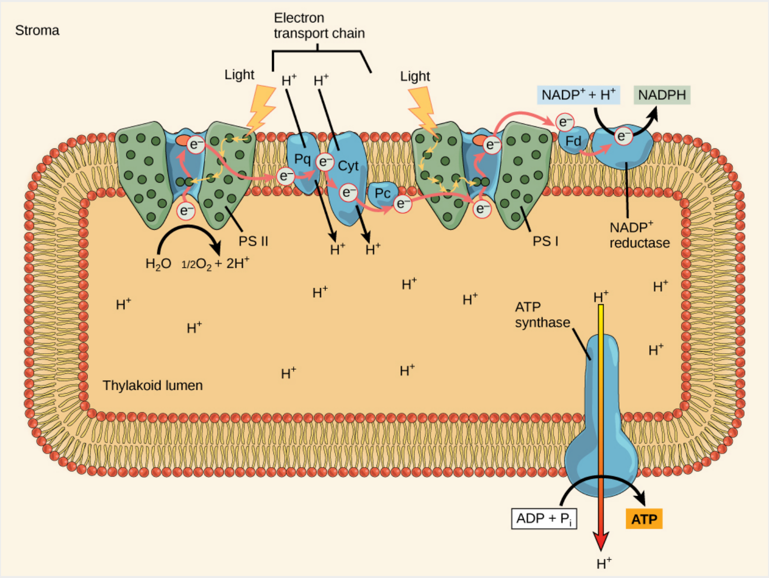
The actual step that converts light energy into chemical energy takes place in a multiprotein

complex called a **photosystem**, two types of which are found embedded in the thylakoid membrane: **photosystem II** (PSII) and **photosystem I** (PSI) ([link]). The two complexes differ on the basis of what they oxidize (that is, the source of the low-energy electron supply) and what they reduce (the place to which they deliver their energized electrons).

Both photosystems have the same basic structure; a number of **antenna proteins** to which the chlorophyll molecules are bound surround the **reaction center** where the photochemistry takes place. Each photosystem is serviced by the **light-harvesting complex**, which passes energy from sunlight to the reaction center; it consists of multiple antenna proteins that contain a mixture of 300 to 400 chlorophyll *a* and *b* molecules as well as other pigments like carotenoids. The absorption of a single **photon** or distinct quantity or “packet” of light by any of the chlorophylls pushes that molecule into an excited state. In short, the light energy has now been captured by biological molecules but is not stored in any useful form yet. The energy is transferred from chlorophyll to chlorophyll until eventually (after about a millionth of a second), it is delivered to the reaction center. Up to this point, only energy has been transferred between molecules, not electrons.

Visual Connection

In the photosystem II (PSII) reaction center, energy from sunlight is used to extract electrons from water. The electrons travel through the chloroplast electron transport chain to photosystem I (PSI), which reduces $NADP^+$ to NADPH. The electron transport chain moves protons across the thylakoid membrane into the lumen. At the same time, splitting of water adds protons to the lumen, and reduction of NADPH removes protons from the stroma. The net result is a low pH in the thylakoid lumen, and a high pH in the stroma. ATP synthase uses this electrochemical gradient to make ATP.



What is the initial source of electrons for the chloroplast electron transport chain?

1. water
2. oxygen

3. carbon dioxide
4. NADPH

The reaction center contains a pair of chlorophyll *a* molecules with a special property. Those two chlorophylls can undergo oxidation upon excitation; they can actually give up an electron in a process called a **photoact**. It is at this step in the reaction center during photosynthesis that light energy is converted into an excited electron. All of the subsequent steps involve getting that electron onto the energy carrier NADPH for delivery to the Calvin cycle where the electron is deposited onto carbon for long-term storage in the form of a carbohydrate. PSII and PSI are two major components of the photosynthetic **electron transport chain**, which also includes the *cytochrome complex*. The cytochrome complex, an enzyme composed of two protein complexes, transfers the electrons from the carrier molecule plastoquinone (Pq) to the protein plastocyanin (Pc), thus enabling both the transfer of protons across the thylakoid membrane and the transfer of electrons from PSII to PSI.

The reaction center of PSII (called **P680**) delivers its high-energy electrons, one at the time, to the **primary electron acceptor**, and through the electron transport chain (Pq to cytochrome complex to plastocyanine) to PSI. P680's missing electron is

replaced by extracting a low-energy electron from water; thus, water is “split” during this stage of photosynthesis, and PSII is re-reduced after every photoact. Splitting one H_2O molecule releases two electrons, two hydrogen atoms, and one atom of oxygen. However, splitting two molecules is required to form one molecule of diatomic O_2 gas. About 10 percent of the oxygen is used by mitochondria in the leaf to support oxidative phosphorylation. The remainder escapes to the atmosphere where it is used by aerobic organisms to support respiration.

As electrons move through the proteins that reside between PSII and PSI, they lose energy. This energy is used to move hydrogen atoms from the stromal side of the membrane to the thylakoid lumen. Those hydrogen atoms, plus the ones produced by splitting water, accumulate in the thylakoid lumen and will be used synthesize ATP in a later step. Because the electrons have lost energy prior to their arrival at PSI, they must be re-energized by PSI, hence, another photon is absorbed by the PSI antenna. That energy is relayed to the PSI reaction center (called **P700**). P700 is oxidized and sends a high-energy electron to $NADP^+$ to form NADPH. Thus, PSII captures the energy to create proton gradients to make ATP, and PSI captures the energy to reduce $NADP^+$ into NADPH. The two photosystems work in concert, in part, to guarantee that the production of NADPH will roughly equal the production of ATP.

Other mechanisms exist to fine-tune that ratio to exactly match the chloroplast's constantly changing energy needs.

Generating an Energy Carrier: ATP

As in the intermembrane space of the mitochondria during cellular respiration, the buildup of hydrogen ions inside the thylakoid lumen creates a *concentration gradient*. The passive diffusion of hydrogen ions from high concentration (in the thylakoid lumen) to low concentration (in the stroma) is harnessed to create ATP, just as in the electron transport chain of cellular respiration. The ions build up energy because of diffusion and because they all have the same electrical charge, repelling each other.

To release this energy, hydrogen ions will rush through any opening, similar to water jetting through a hole in a dam. In the thylakoid, that opening is a passage through a specialized protein channel called the ATP synthase. The energy released by the hydrogen ion stream allows ATP synthase to attach a third phosphate group to ADP, which forms a molecule of ATP ([link]). The flow of hydrogen ions through ATP synthase is called chemiosmosis because the ions move from an area of high to an area of low concentration through a semi-permeable structure of the thylakoid.

Link to Learning

Visit this site and click through the animation to view the process of photosynthesis within a leaf.

Section Summary

The pigments of the first part of photosynthesis, the light-dependent reactions, absorb energy from sunlight. A photon strikes the antenna pigments of photosystem II to initiate photosynthesis. The energy travels to the reaction center that contains chlorophyll *a* and then to the electron transport chain, which pumps hydrogen ions into the thylakoid interior. This action builds up a high concentration of hydrogen ions. The hydrogen ions flow through ATP synthase during chemiosmosis to form molecules of ATP, which are used for the formation of sugar molecules in the second stage of photosynthesis. Photosystem I absorbs a second photon, which results in the formation of an NADPH molecule, another energy and reducing carrier for the light-independent reactions.

Visual Connection Questions

[link] What is the source of electrons for the chloroplast electron transport chain?

1. Water
2. Oxygen
3. Carbon dioxide
4. NADPH

[link] A.

Review Questions

Which of the following structures is *not* a component of a photosystem?

1. ATP synthase
2. antenna molecule
3. reaction center
4. primary electron acceptor

A

How many photons does it take to fully reduce one molecule of $NADP^+$ to NADPH?

1. 1

2. 2
3. 4
4. 8

B

Which complex is *not* involved in the establishment of conditions for ATP synthesis?

1. photosystem I
2. ATP synthase
3. photosystem II
4. cytochrome complex

C

From which component of the light-dependent reactions does NADPH form most directly?

1. photosystem II
2. photosystem I
3. cytochrome complex
4. ATP synthase

B

Three of the same species of plant are each grown under a different colored light for the same amount of time. Plant A is grown under blue light, Plant B is grown under green light, and Plant C is grown under orange light. Assuming the plants use only chlorophyll *a* and chlorophyll *b* for photosynthesis, what would be the predicted order of the plants from most growth to least growth?

1. A, C, B
2. A, B, C
3. C, A, B
4. B, A, C

A

Plants containing only chlorophyll *b* are exposed to radiation with the following wavelengths: 10nm (x-rays), 450nm (blue light), 670nm (red light), and 800nm (infrared light). Which plants harness the most energy for photosynthesis?

1. X-ray irradiated plants
2. Blue light irradiated plants
3. Red light irradiated plants
4. Infrared irradiated plants

B

Critical Thinking Questions

Describe the pathway of electron transfer from photosystem II to photosystem I in light-dependent reactions.

A photon of light hits an antenna molecule in photosystem II, and the energy released by it travels through other antenna molecules to the reaction center. The energy causes an electron to leave a molecule of chlorophyll *a* to a primary electron acceptor protein. The electron travels through the electron transport chain and is accepted by a pigment molecule in photosystem I.

What are the roles of ATP and NADPH in photosynthesis?

Both of these molecules carry energy; in the case of NADPH, it has reducing power that is used to fuel the process of making carbohydrate molecules in light-independent reactions.

How and why would the end products of photosynthesis be changed if a plant had a mutation that eliminated its photosystem II complex?

Knocking out photosystem II would eliminate the production of oxygen and ATP during photosynthesis. Photosystem II splits water into oxygen atoms, hydrogen protons that remain in the thylakoid lumen, and hydrogen-derived electrons that move from the reaction center into the electron transport chain. The transfer of an electron through the electron transport chain provides the energy to pump more protons into the thylakoid lumen to maintain a higher concentration of protons there. Moving protons across the thylakoid membrane back to the stroma provides the energy for ATP synthase to produce ATP. Without this proton gradient, ATP will not be synthesized.

Glossary

absorption spectrum
: range of wavelengths of electromagnetic radiation absorbed by a given substance

antenna protein
: pigment molecule that directly absorbs light and transfers the energy absorbed to other

pigment molecules

carotenoid
: photosynthetic pigment (yellow-orange-red) that functions to dispose of excess energy

chlorophyll *a*
: form of chlorophyll that absorbs violet-blue and red light and consequently has a bluish-green color; the only pigment molecule that performs the photochemistry by getting excited and losing an electron to the electron transport chain

chlorophyll *b*
: accessory pigment that absorbs blue and red-orange light and consequently has a yellowish-green tint

cytochrome complex
: group of reversibly oxidizable and reducible proteins that forms part of the electron transport chain between photosystem II and photosystem I

electromagnetic spectrum
: range of all possible frequencies of radiation

electron transport chain
: group of proteins between PSII and PSI that pass energized electrons and use the energy released by the electrons to move hydrogen

ions against their concentration gradient into the thylakoid lumen

light harvesting complex
: complex that passes energy from sunlight to the reaction center in each photosystem; it consists of multiple antenna proteins that contain a mixture of 300 to 400 chlorophyll *a* and *b* molecules as well as other pigments like carotenoids

P680
: reaction center of photosystem II

P700
: reaction center of photosystem I

photoact
: ejection of an electron from a reaction center using the energy of an absorbed photon

photon
: distinct quantity or “packet” of light energy

photosystem
: group of proteins, chlorophyll, and other pigments that are used in the light-dependent reactions of photosynthesis to absorb light energy and convert it into chemical energy

photosystem I
: integral pigment and protein complex in

thylakoid membranes that uses light energy to transport electrons from plastocyanin to NADP+ (which becomes reduced to NADPH in the process)

photosystem II
: integral protein and pigment complex in thylakoid membranes that transports electrons from water to the electron transport chain; oxygen is a product of PSII

primary electron acceptor
: pigment or other organic molecule in the reaction center that accepts an energized electron from the reaction center

reaction center
: complex of chlorophyll molecules and other organic molecules that is assembled around a special pair of chlorophyll molecules and a primary electron acceptor; capable of undergoing oxidation and reduction

spectrophotometer
: instrument that can measure transmitted light and compute the absorption

wavelength
: distance between consecutive points of equal position (two crests or two troughs) of a wave in a graphic representation; inversely proportional to the energy of the radiation

Using Light Energy to Make Organic Molecules

By the end of this section, you will be able to do the following:

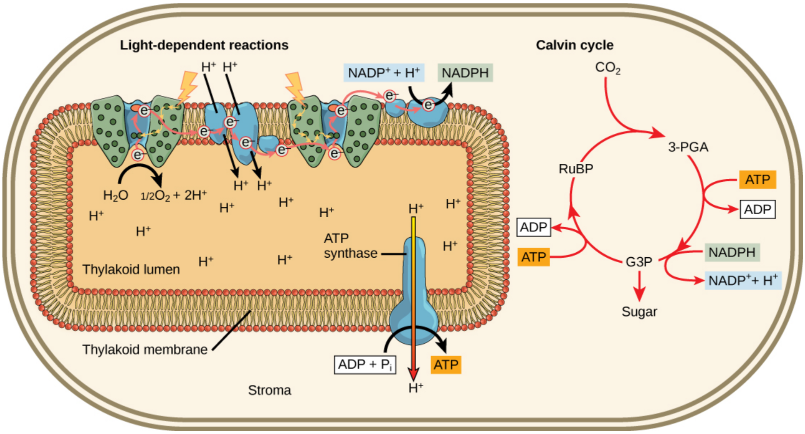
- Describe the Calvin cycle
- Define carbon fixation
- Explain how photosynthesis works in the energy cycle of all living organisms

After the energy from the sun is converted into chemical energy and temporarily stored in ATP and NADPH molecules, the cell has the fuel needed to build carbohydrate molecules for long-term energy storage. The products of the light-dependent reactions, ATP and NADPH, have lifespans in the range of millionths of seconds, whereas the products of the light-independent reactions (carbohydrates and other forms of reduced carbon) can survive almost indefinitely. The carbohydrate molecules made will have a backbone of carbon atoms. But where does the carbon come from? It comes from carbon dioxide—the gas that is a waste product of respiration in microbes, fungi, plants, and animals. Light reactions harness energy from the sun to produce chemical bonds, ATP, and NADPH. These energy-carrying molecules are made in the stroma where carbon fixation takes place. The harsh conditions of the desert have led plants like these cacti to evolve variations of the light-independent reactions of photosynthesis. These variations increase the efficiency of water usage, helping to

conserve water and energy. (credit: Piotr Wojtkowski)

The Calvin Cycle

In plants, carbon dioxide (CO_2) enters the leaves through stomata, where it diffuses over short distances through intercellular spaces until it reaches the mesophyll cells. Once in the mesophyll cells, CO_2 diffuses into the stroma of the chloroplast—the site of light-independent reactions of photosynthesis. These reactions actually have several names associated with them. Another term, the **Calvin cycle**, is named for the man who discovered it, and because these reactions function as a cycle. Others call it the Calvin-Benson cycle to include the name of another scientist involved in its discovery. The most outdated name is “dark reaction,” because light is not directly required ([link]). However, the term dark reaction can be misleading because it implies incorrectly that the reaction only occurs at night or is independent of light, which is why most scientists and instructors no longer use it.



The light-independent reactions of the Calvin cycle can be organized into three basic stages: *fixation, reduction, and regeneration.*

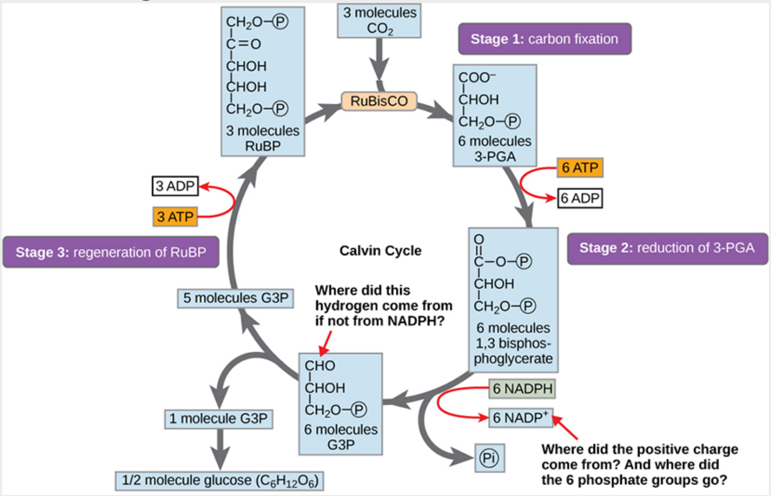
Stage 1: Fixation

In the stroma, in addition to CO_2, two other components are present to initiate the light-independent reactions: an enzyme called ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), and three molecules of ribulose bisphosphate (RuBP), as shown in [link]. RuBP has five atoms of carbon, flanked by two phosphates.

Visual Connection

The Calvin cycle has three stages. In stage 1, the enzyme RuBisCO incorporates carbon dioxide into an organic molecule, 3-PGA. In stage 2, the organic

molecule is reduced using electrons supplied by NADPH. In stage 3, RuBP, the molecule that starts the cycle, is regenerated so that the cycle can continue. Only one carbon dioxide molecule is incorporated at a time, so the cycle must be completed three times to produce a single three-carbon G3P molecule, and six times to produce a six-carbon glucose molecule.



Which of the following statements is true?

1. In photosynthesis, oxygen, carbon dioxide, ATP, and NADPH are reactants. G3P and water are products.
2. In photosynthesis, chlorophyll, water, and carbon dioxide are reactants. G3P and oxygen are products.
3. In photosynthesis, water, carbon dioxide, ATP, and NADPH are reactants. RuBP and oxygen are products.
4. In photosynthesis, water and carbon dioxide

are reactants. G3P and oxygen are products.

RuBisCO catalyzes a reaction between CO_2 and RuBP. For each CO_2 molecule that reacts with one RuBP, two molecules of another compound 3-phospho glyceric acid (3-PGA) form. PGA has three carbons and one phosphate. Each turn of the cycle involves only one RuBP and one carbon dioxide and forms two molecules of 3-PGA. The number of carbon atoms remains the same, as the atoms move to form new bonds during the reactions (3 C atoms from $3CO_2$ + 15 C atoms from 3RuBP = 18 C atoms in 6 molecules of 3-PGA). This process is called **carbon fixation**, because CO_2 is "fixed" from an inorganic form into organic molecules.

Stage 2: Reduction

ATP and NADPH are used to convert the six molecules of 3-PGA into six molecules of a chemical called glyceraldehyde 3-phosphate (G3P). That is a reduction reaction because it involves the gain of electrons by 3-PGA. (Recall that a **reduction** is the gain of an electron by an atom or molecule.) Six molecules of both ATP and NADPH are used. For ATP, energy is released with the loss of the terminal phosphate atom, converting it into ADP; for NADPH, both energy and a hydrogen atom are lost, converting it into $NADP^+$. Both of these molecules

return to the nearby light-dependent reactions to be reused and re-energized.

Stage 3: Regeneration

Interestingly, at this point, only one of the G3P molecules leaves the Calvin cycle and is sent to the cytoplasm to contribute to the formation of other compounds needed by the plant. Because the G3P exported from the chloroplast has three carbon atoms, it takes three "turns" of the Calvin cycle to fix enough net carbon to export one G3P. But each turn makes two G3Ps, thus three turns make six G3Ps. One is exported while the remaining five G3P molecules remain in the cycle and are used to regenerate RuBP, which enables the system to prepare for more CO_2 to be fixed. Three more molecules of ATP are used in these regeneration reactions.

Link to Learning

This link leads to an animation of photosynthesis and the Calvin cycle.

Evolution Connection

Photosynthesis

During the evolution of photosynthesis, a major

shift occurred from the bacterial type of photosynthesis that involves only one photosystem and is typically anoxygenic (does not generate oxygen) into modern oxygenic (does generate oxygen) photosynthesis, employing two photosystems. This modern oxygenic photosynthesis is used by many organisms—from giant tropical leaves in the rainforest to tiny cyanobacterial cells—and the process and components of this photosynthesis remain largely the same. Photosystems absorb light and use electron transport chains to convert energy into the chemical energy of ATP and NADH. The subsequent light-independent reactions then assemble carbohydrate molecules with this energy. In the harsh dry heat of the desert, plants must conserve every drop of water must be used to survive. Because stomata must open to allow for the uptake of CO_2, water escapes from the leaf during active photosynthesis. Desert plants have evolved processes to conserve water and deal with harsh conditions. Mechanisms to capture and store CO_2 allows plants to adapt to living with less water. Some plants such as cacti ([link]) can prepare materials for photosynthesis during the night by a temporary carbon fixation/storage process, because opening the stomata at this time conserves water due to cooler temperatures. During the day cacti use the captured CO_2 for photosynthesis, and keep their stomata closed.


Photosynthesis consumes carbon dioxide and produces oxygen. Aerobic respiration consumes oxygen and produces carbon dioxide. These two processes play an important role in the carbon cycle. (credit: modification of work by Stuart Bassil)

The Energy Cycle

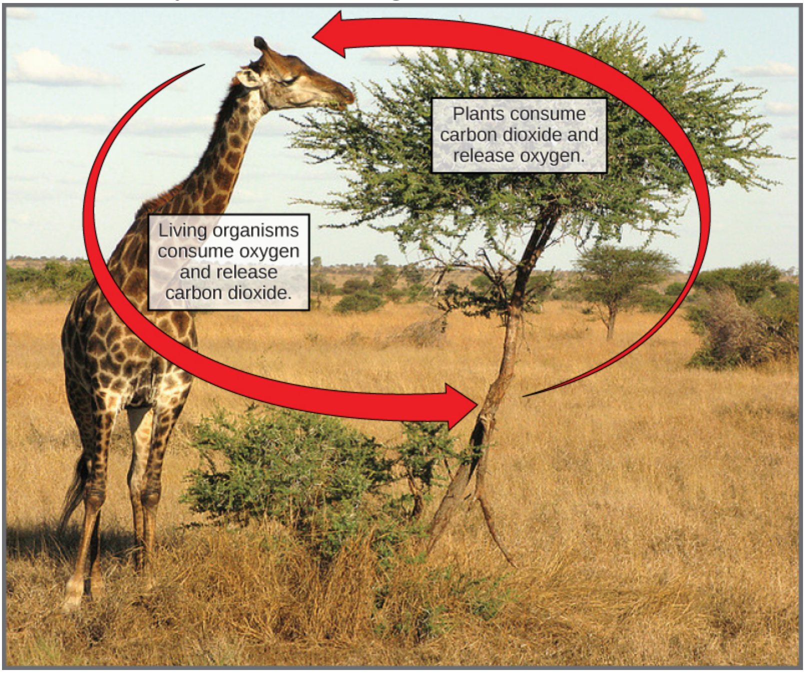
Whether the organism is a bacterium, plant, or animal, all living things access energy by breaking down carbohydrate and other carbon-rich organic

molecules. But if plants make carbohydrate molecules, why would they need to break them down, especially when it has been shown that the gas organisms release as a “waste product” (CO_2) acts as a substrate for the formation of more food in photosynthesis? Remember, living things need energy to perform life functions. In addition, an organism can either make its own food or eat another organism—either way, the food still needs to be broken down. Finally, in the process of breaking down food, called cellular respiration, heterotrophs release needed energy and produce “waste” in the form of CO_2 gas.

However, in nature, there is no such thing as “waste.” Every single atom of matter and energy is conserved, recycled over and over infinitely. Substances change form or move from one type of molecule to another, but their constituent atoms never disappear ([link]).

In reality, CO_2 is no more a form of waste than oxygen is wasteful to photosynthesis. Both are byproducts of reactions that move on to other reactions. Photosynthesis absorbs light energy to build carbohydrates in chloroplasts, and aerobic cellular respiration releases energy by using oxygen to metabolize carbohydrates in the cytoplasm and mitochondria. Both processes use electron transport chains to capture the energy necessary to drive other reactions. These two powerhouse processes,

photosynthesis and cellular respiration, function in biological, cyclical harmony to allow organisms to access life-sustaining energy that originates millions of miles away in a burning star humans call the sun.



Section Summary

Using the energy carriers formed in the first steps of photosynthesis, the light-independent reactions, or the Calvin cycle, take in CO_2 from the atmosphere. An enzyme, RuBisCO, catalyzes a reaction with CO_2 and another organic compound, RuBP. After three cycles, a three-carbon molecule of G3P leaves the cycle to become part of a carbohydrate molecule. The remaining G3P molecules stay in the cycle to be

regenerated into RuBP, which is then ready to react with more CO_2. Photosynthesis forms an energy cycle with the process of cellular respiration. Because plants contain both chloroplasts and mitochondria, they rely upon both photosynthesis and respiration for their ability to function in both the light and dark, and to be able to interconvert essential metabolites.

Visual Connection Questions

[link] Which of the following statements is true?

1. In photosynthesis, oxygen, carbon dioxide, ATP, and NADPH are reactants. G3P and water are products.
2. In photosynthesis, chlorophyll, water, and carbon dioxide are reactants. G3P and oxygen are products.
3. In photosynthesis, water, carbon dioxide, ATP, and NADPH are reactants. RuBP and oxygen are products.
4. In photosynthesis, water and carbon dioxide are reactants. G3P and oxygen are products.

[link] D

Review Questions

Which molecule must enter the Calvin cycle continually for the light-independent reactions to take place?

1. RuBisCO
2. RuBP
3. 3-PGA
4. CO_2

D

Which order of molecular conversions is correct for the Calvin cycle?

1. RuBP + G3P → 3-PGA → sugar
2. RuBisCO → CO_2 → RuBP → G3P
3. RuBP + CO_2 → [RuBisCO] 3-PGA → G3P
4. CO_2 → 3-PGA → RuBP → G3P

C

Where in eukaryotic cells does the Calvin cycle take place?

1. thylakoid membrane
2. thylakoid lumen
3. chloroplast stroma
4. granum

C

Which statement correctly describes carbon fixation?

1. the conversion of CO_2 into an organic compound
2. the use of RuBisCO to form 3-PGA
3. the production of carbohydrate molecules from G3P
4. the formation of RuBP from G3P molecules
5. the use of ATP and NADPH to reduce CO_2

A

If four molecules of carbon dioxide enter the Calvin cycle (four “turns” of the cycle), how many G3P molecules are produced and how many are exported?

1. 4 G3P made, 1 G3P exported
2. 4 G3P made, 2 G3P exported
3. 8 G3P made, 1 G3P exported
4. 8 G3P made, 4 G3P exported

C

Critical Thinking Questions

Why is the third stage of the Calvin cycle called the regeneration stage?

Because RuBP, the molecule needed at the start of the cycle, is regenerated from G3P.

Which part of the light-independent reactions would be affected if a cell could not produce the enzyme RuBisCO?

None of the cycle could take place, because RuBisCO is essential in fixing carbon dioxide. Specifically, RuBisCO catalyzes the reaction between carbon dioxide and RuBP at the start of the cycle.

Why does it take three turns of the Calvin cycle to produce G3P, the initial product of photosynthesis?

Because G3P has three carbon atoms, and each turn of the cycle takes in one carbon atom in the form of carbon dioxide.

Imagine a sealed terrarium containing a plant and a beetle. How does each organism provide resources for the other? Could each organism survive if it was the only living thing in the terrarium? Why or why not?

An energy cycle between a plant and a beetle would be as follows:

1. Plant consumes carbon dioxide and releases oxygen as a byproduct of photosynthesis
2. Beetle consumes oxygen and releases carbon dioxide to create chemical energy during aerobic respiration
3. Plant takes up carbon dioxide from the air
4. Repeat cycle

The plant would also provide a carbon-based food source for the beetle.

1. The beetle is a heterotroph, and would not survive without the plant because it would deplete all the oxygen within the terrarium.
2. The plant is an autotroph and could survive without the beetle, but it would be unlikely to grow. Through photosynthesis, the plant can make and store its own energy in carbon-based molecules, and produce oxygen. The oxygen can then be used to power aerobic respiration in the plant, which releases carbon dioxide. However, since the plant essentially continues to reuse its own resources cycling between carbon- and oxygen-consuming pathways, its growth would be limited.

Compare the flow of energy with the flow of nutrients in a closed, sunny ecosystem consisting of a giraffe and a tree.

In the defined ecosystem, energy would radiate from the Sun, and be absorbed by the chlorophyll in the leaves of the tree. Photosynthesis would occur in the leaves, transforming the light energy into stored chemical energy in the covalent bonds of carbon molecules. The giraffe would eat the

leaves of the tree, and digest the carbon molecules to release energy.

In the same ecosystem, nutrients would cycle between the tree and the giraffe. The giraffe would consume oxygen and release carbon dioxide as its cells perform aerobic respiration to create chemical energy. The tree will consume the released carbon dioxide during photosynthesis to create its own stored chemical energy, and release oxygen as a by-product.

Glossary

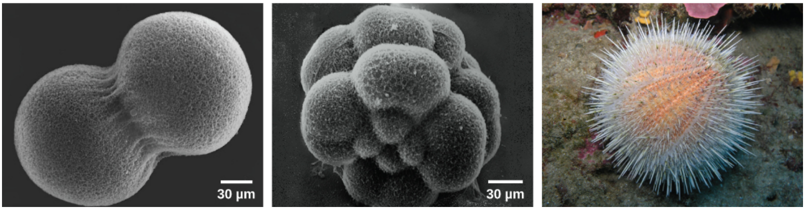
Calvin cycle
: light-independent reactions of photosynthesis that convert carbon dioxide from the atmosphere into carbohydrates using the energy and reducing power of ATP and NADPH

carbon fixation
: process of converting inorganic CO_2 gas into organic compounds

reduction
: gain of electron(s) by an atom or molecule

Introduction

class="introduction" A sea urchin begins life as a single diploid cell (zygote) that (a) divides through cell division to form two genetically identical daughter cells, visible here through scanning electron microscopy (SEM). After four rounds of cell division, (b) there are 16 cells, as seen in this SEM image. After many rounds of cell division, the individual develops into a complex, multicellular organism, as seen in this (c) mature sea urchin. (credit a: modification of work by Evelyn Spiegel, Louisa Howard; credit b: modification of work by Evelyn Spiegel, Louisa Howard; credit c: modification of work by Marco Busdraghi; scale-bar data from Matt Russell)



A human, like every sexually reproducing organism, begins life as a fertilized egg (embryo) or **zygote**. In our species, billions of cell divisions subsequently must occur in a controlled manner in order to produce a complex, multicellular human comprising trillions of cells. Thus, the original single-celled zygote is literally the ancestor of all cells in the body. However, once a human is fully grown, cell reproduction is still necessary to repair and regenerate tissues, and sometimes to increase our

size! In fact, all multicellular organisms use cell division for growth and the maintenance and repair of cells and tissues. Cell division is closely regulated, and the occasional failure of this regulation can have life-threatening consequences. Single-celled organisms may also use cell division as their method of reproduction.

Cell Division

By the end of this section, you will be able to do the following:

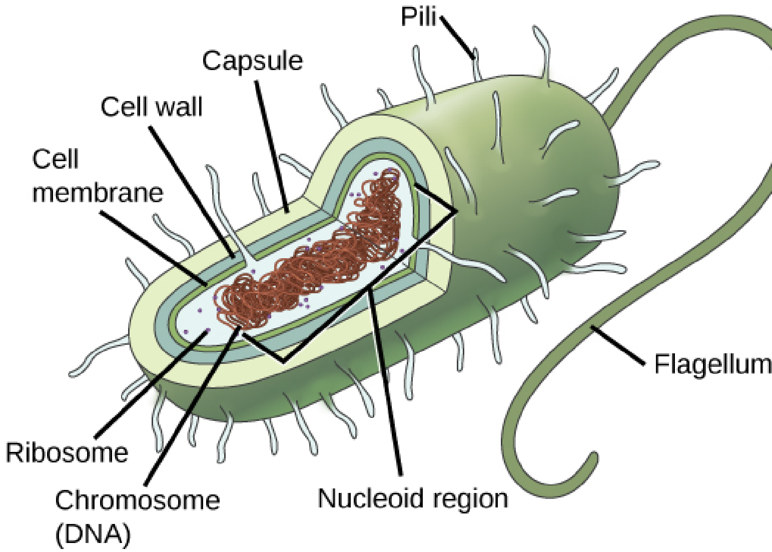
- Describe the structure of prokaryotic and eukaryotic genomes
- Distinguish between chromosomes, genes, and traits
- Describe the mechanisms of chromosome compaction

The continuity of life from one cell to another has its foundation in the reproduction of cells by way of the cell cycle. The cell cycle is an orderly sequence of events that describes the stages of a cell’s life from the division of a single parent cell to the production of two new genetically identical daughter cells.
Prokaryotes, including both Bacteria and Archaea, have a single, circular chromosome located in a central region called the nucleoid. There are 23 pairs of homologous chromosomes in a female human somatic cell. The condensed chromosomes are viewed within the nucleus (top), removed from a cell during mitosis (also called karyokinesis or nuclear division) and spread out on a slide (right), and artificially arranged according to length (left); an arrangement like this is called a karyotype. In this image, the chromosomes were exposed to fluorescent stains for differentiation of the different chromosomes. A method of staining called

"chromosome painting" employs fluorescent dyes that highlight chromosomes in different colors. (credit: National Human Genome Project/NIH)

Genomic DNA

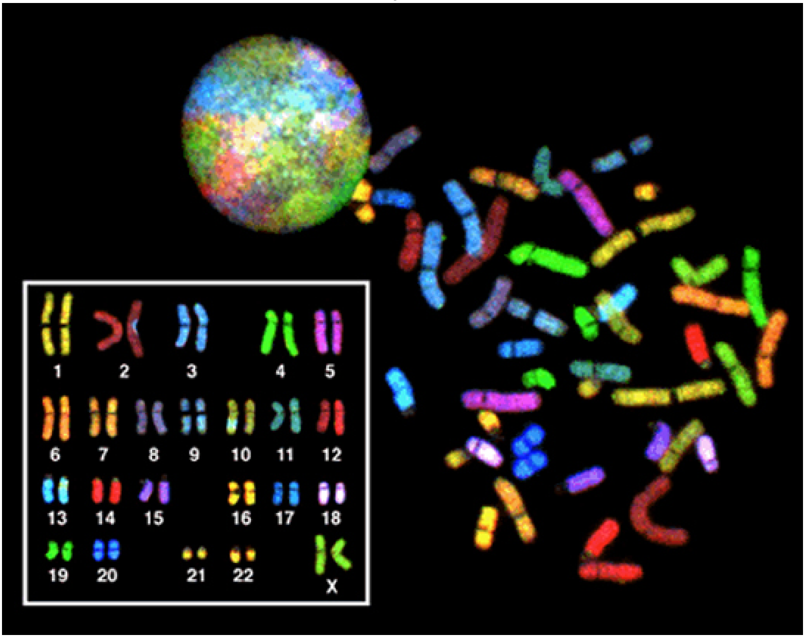
Before discussing the steps a cell must undertake to replicate and divide its DNA, a deeper understanding of the structure and function of a cell's genetic information is necessary. A cell's DNA, packaged as a double-stranded DNA molecule, is called its **genome**. In prokaryotes, the genome is composed of a single, double-stranded DNA molecule in the form of a loop or circle ([link]). The region in the cell containing this genetic material is called a nucleoid. Some prokaryotes also have smaller loops of DNA called plasmids that are not essential for normal growth. Bacteria can exchange these plasmids with other bacteria, sometimes receiving beneficial new genes that the recipient can add to their chromosomal DNA. *Antibiotic resistance* is one trait that often spreads through a bacterial colony through plasmid exchange from resistant donors to recipient cells.



In eukaryotes, the genome consists of several double-stranded linear DNA molecules ([link]). Each species of eukaryotes has a characteristic number of chromosomes in the nuclei of its cells. Human body (somatic) cells have 46 chromosomes, while human **gametes** (sperm or eggs) have 23 chromosomes each. A typical body cell contains two matched or homologous sets of chromosomes (one set from each biological parent)—a configuration known as **diploid**. (Note: The letter *n* is used to represent a single set of chromosomes; therefore, a diploid organism is designated 2*n*.) Human cells that contain one set of chromosomes are called gametes, or sex cells; these are eggs and sperm, and are designated *1n*, or **haploid**.

Upon fertilization, each gamete contributes one set of chromosomes, creating a diploid cell containing

matched pairs of chromosomes called **homologous** ("same knowledge") **chromosomes**. Homologous chromosomes are the same length and have specific nucleotide segments called **genes** in exactly the same location, or **locus**. Genes, the functional units of chromosomes, determine specific characteristics by coding for specific proteins. Traits are the variations of those characteristics. For example, hair color is a characteristic with traits that are blonde, brown, or black, and many colors in between.



Each copy of a homologous pair of chromosomes originates from a different parent; therefore, the different genes (alleles) themselves are not identical, although they code for the same traits such as "hair color." The variation of individuals within a species is due to the specific combination of the genes

inherited from both parents. Even a slightly altered sequence of nucleotides within a gene can result in an alternative trait. For example, there are three possible gene sequences on the human chromosome that code for blood type: sequence A, sequence B, and sequence O. Because all diploid human cells have two copies of the chromosome that determines blood type, the blood type (the trait) is determined by the two alleles of the marker gene that are inherited. It is possible to have two copies of the same gene sequence on both homologous chromosomes, with one on each (for example, AA, BB, or OO), or two different sequences, such as AB, AO, or BO.

Apparently minor variations of traits, such as blood type, eye color, and handedness, contribute to the natural variation found within a species, but even though they seem minor, these traits may be connected with the expression of other traits as of yet unknown. However, if the entire DNA sequence from any pair of human homologous chromosomes is compared, the difference is much less than one percent. The sex chromosomes, X and Y, are the single exception to the rule of homologous chromosome uniformity: Other than a small amount of homology that is necessary to accurately produce gametes, the genes found on the X and Y chromosomes are different.

Double-stranded DNA wraps around histone proteins to form nucleosomes that create the

appearance of “beads on a string.” The nucleosomes are coiled into a 30-nm chromatin fiber. When a cell undergoes mitosis, the chromosomes condense even further.

Eukaryotic Chromosomal Structure and Compaction

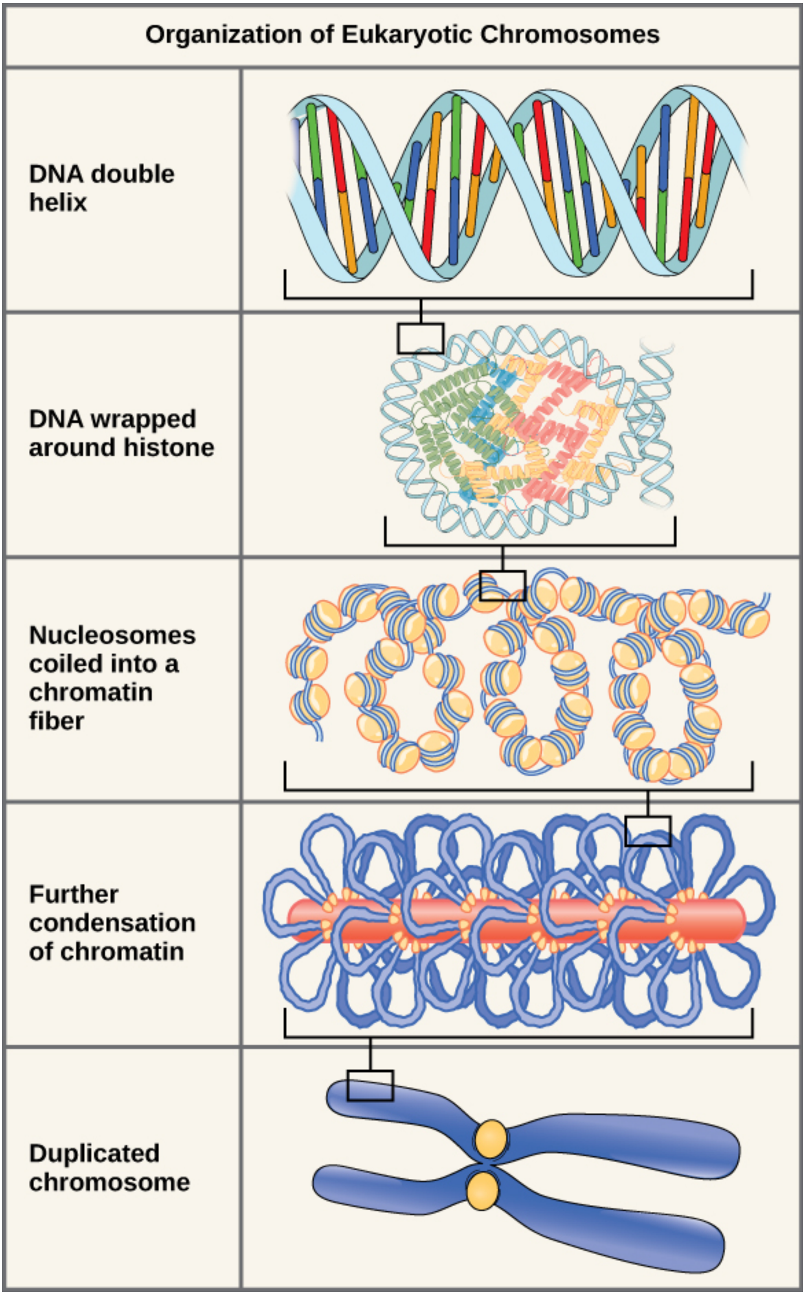
If the DNA from all 46 chromosomes in a human cell nucleus were laid out end-to-end, it would measure approximately two meters; however, its diameter would be only 2 nm! Considering that the size of a typical human cell is about 10 µm (100,000 cells lined up to equal one meter), DNA must be *tightly packaged* to fit in the cell’s nucleus. At the same time, it must also be readily accessible for the genes to be expressed. For this reason, the long strands of DNA are condensed into compact chromosomes during certain stages of the cell cycle. There are a number of ways that chromosomes are compacted.

In the first level of compaction, short stretches of the DNA double helix wrap around a core of eight **histone proteins** at regular intervals along the entire length of the chromosome ([link]). The DNA-histone complex is called chromatin. The beadlike, histone DNA complex is called a **nucleosome**, and DNA connecting the nucleosomes is called linker DNA. A DNA molecule in this form is about seven times shorter than the double helix without the

histones, and the beads are about 10 nm in diameter, in contrast with the 2-nm diameter of a DNA double helix.

The second level of compaction occurs as the nucleosomes and the linker DNA between them coil into a 30-nm chromatin fiber. This coiling further *condenses* the chromosome so that it is now about 50 times shorter than the extended form.

In the third level of compaction, a variety of *fibrous proteins* is used to “pack the chromatin.” These fibrous proteins also ensure that each chromosome in a non-dividing cell occupies a particular area of the nucleus that does not overlap with that of any other chromosome (see the top image in [link]).



DNA replicates in the S phase of interphase, which technically is not a part of mitosis, but must always

precede it. After replication, the chromosomes are composed of two linked sister **chromatids**. When fully compact, the pairs of identically packed chromosomes are bound to each other by cohesin proteins. The connection between the sister chromatids is closest in a region called the **centromere**. The conjoined sister chromatids, with a diameter of about 1 μm, are visible under a light microscope. The centromeric region is highly condensed and thus will appear as a constricted area.

Link to Learning

This animation illustrates the different levels of chromosome packing.
https://www.openstax.org/l/Packaged_DNA

Section Summary

Prokaryotes have a single circular chromosome composed of double-stranded DNA, whereas eukaryotes have multiple, linear chromosomes composed of chromatin wrapped around histones, all of which are surrounded by a nuclear membrane. The 46 chromosomes of human somatic cells are

composed of 22 pairs of autosomes (matched pairs) and a pair of sex chromosomes, which may or may not be matched. This is the 2*n* or diploid state. Human gametes have 23 chromosomes, or one complete set of chromosomes; a set of chromosomes is complete with either one of the sex chromosomes, X or Y. This is the *n* or haploid state. Genes are segments of DNA that code for a specific functional molecule (a protein or RNA). An organism's traits are determined by the genes inherited from each parent. Duplicated chromosomes are composed of two sister chromatids. Chromosomes are compacted using a variety of mechanisms during certain stages of the cell cycle. Several classes of protein are involved in the organization and packing of the chromosomal DNA into a highly condensed structure. The condensing complex compacts chromosomes, and the resulting condensed structure is necessary for chromosomal segregation during mitosis.

Review Questions

A diploid cell has_____ the number of chromosomes as a haploid cell.

1. one-fourth
2. half

3. twice
4. four times

C

An organism's traits are determined by the specific combination of inherited ____.

1. cells.
2. genes.
3. proteins.
4. chromatids.

B

The first level of DNA organization in a eukaryotic cell is maintained by which molecule?

1. cohesin
2. condensin
3. chromatin
4. histone

D

Identical copies of chromatin held together by cohesin at the centromere are called ____.

1. histones.
2. nucleosomes.
3. chromatin.
4. sister chromatids.

D

Critical Thinking Questions

Compare and contrast a human somatic cell to a human gamete.

Human somatic cells have 46 chromosomes: 22 pairs and 2 sex chromosomes that may or may not form a pair. This is the 2*n* or diploid condition. Human gametes have 23 chromosomes, one each of 23 unique chromosomes, one of which is a sex chromosome. This is the *n* or haploid condition.

What is the relationship between a genome, chromosomes, and genes?

The genome consists of the sum total of an organism's chromosomes. Each chromosome contains hundreds and sometimes thousands of genes, segments of DNA that code for a polypeptide or RNA, and a large amount of DNA with no known function.

Eukaryotic chromosomes are thousands of times longer than a typical cell. Explain how chromosomes can fit inside a eukaryotic nucleus.

The DNA double helix is wrapped around histone proteins to form structures called nucleosomes. Nucleosomes and the linker DNA in between them are coiled into a 30-nm fiber. During cell division, chromatin is further condensed by packing proteins.

Glossary

cell cycle
: ordered sequence of events through which a cell passes between one cell division and the next

centromere
: region at which sister chromatids are bound

together; a constricted area in condensed chromosomes

chromatid
: single DNA molecule of two strands of duplicated DNA and associated proteins held together at the centromere

diploid
: cell, nucleus, or organism containing two sets of chromosomes (2*n*)

gamete
: haploid reproductive cell or sex cell (sperm, pollen grain, or egg)

gene
: physical and functional unit of heredity, a sequence of DNA that codes for a protein.

genome
: total genetic information of a cell or organism

haploid
: cell, nucleus, or organism containing one set of chromosomes (*n*)

histone
: one of several similar, highly conserved, low molecular weight, basic proteins found in the chromatin of all eukaryotic cells; associates with DNA to form nucleosomes

homologous chromosomes
: chromosomes of the same morphology with genes in the same location; diploid organisms have pairs of homologous chromosomes (homologs), with each homolog derived from a different parent

locus
: position of a gene on a chromosome

nucleosome
: subunit of chromatin composed of a short length of DNA wrapped around a core of histone proteins

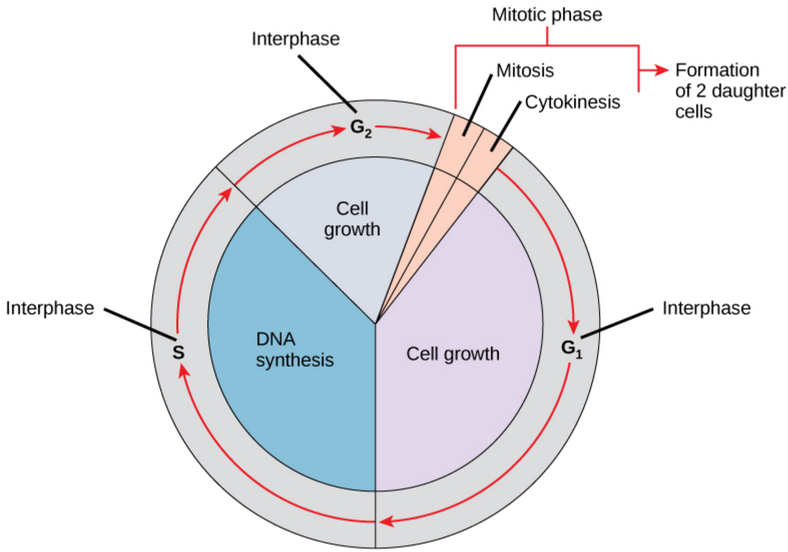
The Cell Cycle

By the end of this section, you will be able to do the following:

- Describe the three stages of interphase
- Discuss the behavior of chromosomes during karyokinesis/mitosis
- Explain how the cytoplasmic content is divided during cytokinesis
- Define the quiescent G_0 phase

The **cell cycle** is an ordered series of events involving cell growth and cell division that produces two new daughter cells. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages of growth, DNA replication, and nuclear and cytoplasmic division that ultimately produces two identical (clone) cells. The cell cycle has two major phases: interphase and the mitotic phase ([link]). During **interphase**, the cell grows and DNA is replicated. During the **mitotic phase**, the replicated DNA and cytoplasmic contents are separated, and the cell cytoplasm is typically partitioned by a third process of the cell cycle called **cytokinesis**. We should note, however, that interphase and mitosis (kayrokinesis) may take place without cytokinesis, in which case cells with multiple nuclei (multinucleate cells) are produced. The cell cycle in multicellular organisms consists of interphase and the mitotic phase. During interphase, the cell grows and the nuclear DNA is duplicated.

Interphase is followed by the mitotic phase. During the mitotic phase, the duplicated chromosomes are segregated and distributed into daughter nuclei. Following mitosis, the cytoplasm is usually divided as well by cytokinesis, resulting in two genetically identical daughter cells.



Interphase

During interphase, the cell undergoes normal growth processes while also preparing for cell division. In order for a cell to move from interphase into the mitotic phase, many internal and external conditions must be met. The three stages of interphase are called *G1, S, and G2.*

G1 Phase (First Gap)

The first stage of interphase is called the **G_1 phase** (first gap) because, from a microscopic point of view, little change is visible. However, during the G_1 stage, the cell is quite active at the biochemical level. The cell is accumulating the building blocks of chromosomal DNA and the associated proteins as well as accumulating sufficient energy reserves to complete the task of replicating each chromosome in the nucleus.

S Phase (Synthesis of DNA)

Throughout interphase, nuclear DNA remains in a semi-condensed chromatin configuration. In the **S phase**, DNA replication can proceed through the mechanisms that result in the formation of identical pairs of DNA molecules—sister chromatids—that are firmly attached to the centromeric region. The centrosome is also duplicated during the S phase. The two centrosomes of homologous chromosomes will give rise to the **mitotic spindle**, the apparatus that orchestrates the movement of chromosomes during mitosis. For example, roughly at the center of each animal cell, the centrosomes are associated with a pair of rod-like objects, the **centrioles**, which are positioned at right angles to each other. Centrioles help organize cell division. We should note, however, that centrioles are not present in the centrosomes of other eukaryotic organisms, such as plants and most fungi.

G2 Phase (Second Gap)

In the **G2 phase**, the cell replenishes its energy stores and synthesizes proteins necessary for chromosome manipulation and movement. Some cell organelles are duplicated, and the cytoskeleton is dismantled to provide resources for the mitotic phase. There may be additional cell growth during G2. The final preparations for the mitotic phase must be completed before the cell is able to enter the first stage of mitosis.
During prometaphase, mitotic spindle microtubules from opposite poles attach to each sister chromatid at the kinetochore. In anaphase, the connection between the sister chromatids breaks down, and the microtubules pull the chromosomes toward opposite poles.During cytokinesis in animal cells, a ring of actin filaments forms at the metaphase plate. The ring contracts, forming a cleavage furrow, which divides the cell in two. In plant cells, Golgi vesicles coalesce at the former metaphase plate, forming a phragmoplast. A cell plate formed by the fusion of the vesicles of the phragmoplast grows from the center toward the cell walls, and the membranes of the vesicles fuse to form a plasma membrane that divides the cell in two.

The Mitotic Phase

The mitotic phase is a multistep process during which the duplicated chromosomes are aligned,

separated, and move into two new, identical daughter cells. The first portion of the mitotic phase is called **karyokinesis**, or nuclear division. As we have just seen, the second portion of the mitotic phase (and often viewed as a process separate from and following mitosis) is called cytokinesis—the physical separation of the cytoplasmic components into the two daughter cells.

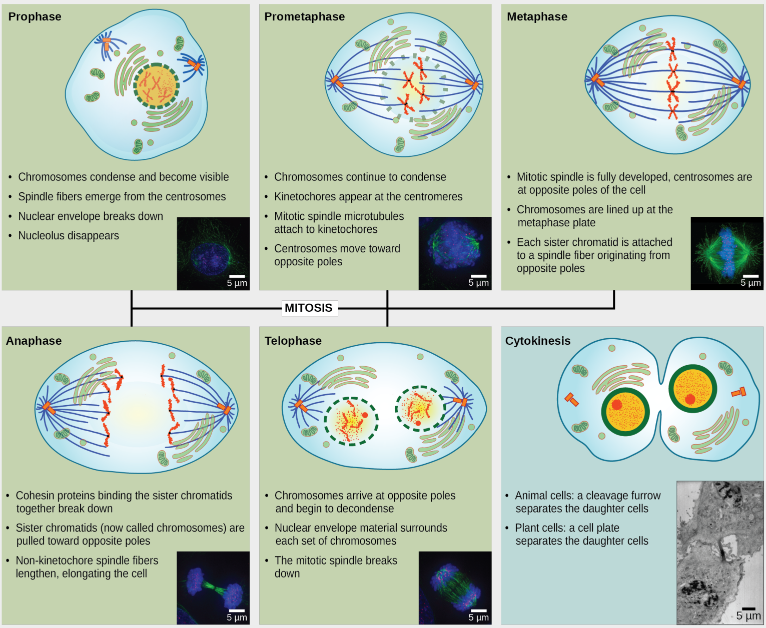
Link to Learning
Revisit the stages of mitosis at this site.

Karyokinesis (Mitosis)

Karyokinesis, also known as **mitosis**, is divided into a series of phases—prophase, prometaphase, metaphase, anaphase, and telophase—that result in the division of the cell nucleus ([link]).

Visual Connection
Karyokinesis (or mitosis) is divided into five stages—prophase, prometaphase, metaphase, anaphase, and telophase. The pictures at the bottom were taken by fluorescence microscopy (hence, the black background) of cells artificially stained by

fluorescent dyes: blue fluorescence indicates DNA (chromosomes) and green fluorescence indicates microtubules (spindle apparatus). (credit “mitosis drawings”: modification of work by Mariana Ruiz Villareal; credit “micrographs”: modification of work by Roy van Heesbeen; credit “cytokinesis micrograph”: Wadsworth Center/New York State Department of Health; scale-bar data from Matt Russell)



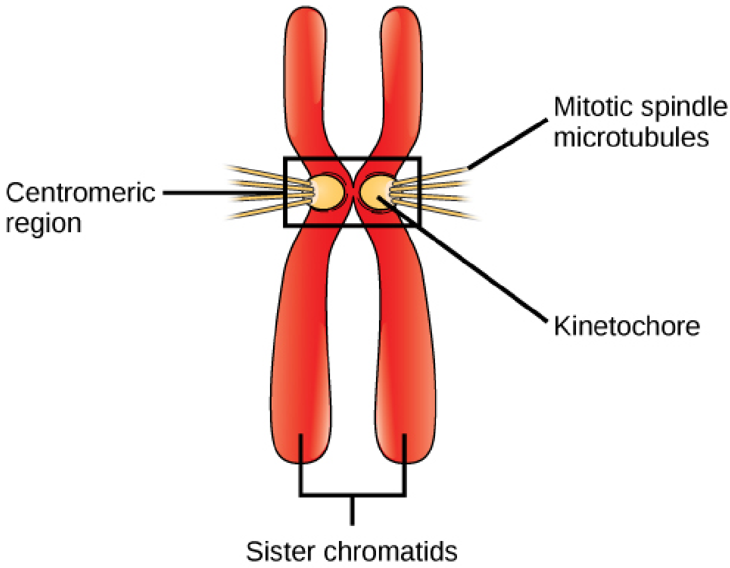
Which of the following is the correct order of events in mitosis?

1. Sister chromatids line up at the metaphase plate. The kinetochore becomes attached to the mitotic spindle. The nucleus reforms and the cell divides. Cohesin proteins break down and the sister chromatids separate.

2. The kinetochore becomes attached to the mitotic spindle. Cohesin proteins break down and the sister chromatids separate. Sister chromatids line up at the metaphase plate. The nucleus reforms and the cell divides.
3. The kinetochore becomes attached to the cohesin proteins. Sister chromatids line up at the metaphase plate. The kinetochore breaks down and the sister chromatids separate. The nucleus reforms and the cell divides.
4. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

Prophase (the "first phase"): the nuclear envelope starts to dissociate into small vesicles, and the membranous organelles (such as the Golgi complex [Golgi apparatus] and the endoplasmic reticulum), fragment and disperse toward the periphery of the cell. The nucleolus disappears (disperses) as well, and the centrosomes begin to move to opposite poles of the cell. Microtubules that will form the *mitotic spindle* extend between the centrosomes, *pushing them farther apart* as the microtubule fibers lengthen. The sister chromatids begin to coil more tightly with the aid of **condensin proteins** and now become visible under a light microscope.

Prometaphase (the “first change phase”): Many processes that began in prophase continue to advance. The remnants of the nuclear envelope fragment further, and the mitotic spindle continues to develop as more microtubules assemble and stretch across the length of the former nuclear area. Chromosomes become even more condensed and discrete. Each sister chromatid develops a protein structure called a **kinetochore** in its centromeric region ([link]). The proteins of the kinetochore attract and bind to the mitotic spindle microtubules. As the spindle microtubules extend from the centrosomes, some of these microtubules come into contact with and firmly bind to the kinetochores. Once a mitotic fiber attaches to a chromosome, the chromosome will be oriented until the kinetochores of sister chromatids face the *opposite poles*. Eventually, all the sister chromatids will be attached via their kinetochores to microtubules from opposing poles. Spindle microtubules that do not engage the chromosomes are called **polar microtubules**. These microtubules overlap each other midway between the two poles and contribute to *cell elongation*. Astral microtubules are located near the poles, aid in spindle orientation, and are required for the regulation of mitosis.



Metaphase (the “change phase”): All the chromosomes are aligned in a plane called the **metaphase plate**, or the equatorial plane, roughly midway between the two poles of the cell. The sister chromatids are still tightly attached to each other by cohesin proteins. At this time, the chromosomes are maximally condensed.

Anaphase (“upward phase”): The cohesin proteins degrade, and the sister chromatids separate at the centromere. Each chromatid, now called a single chromosome, is pulled rapidly toward the centrosome to which its microtubule is attached. The cell becomes visibly elongated (oval shaped) as the polar microtubules slide against each other at the metaphase plate where they overlap.

Telophase (the "distance phase"): the chromosomes reach the opposite poles and begin to *decondense* (unravel), relaxing once again into a stretched-out chromatin configuration. The mitotic spindles are depolymerized into tubulin monomers that will be used to assemble cytoskeletal components for each daughter cell. Nuclear envelopes form around the chromosomes, and nucleosomes appear within the nuclear area.

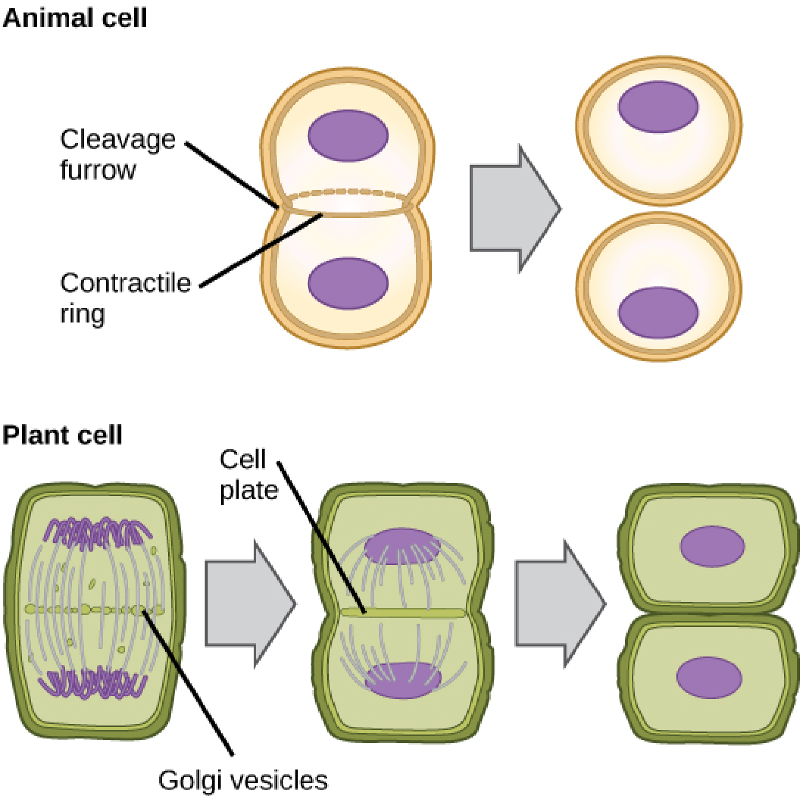
Cytokinesis

Cytokinesis, or "cell motion," is sometimes viewed as the second main stage of the mitotic phase, during which cell division is completed via the physical separation of the cytoplasmic components into two daughter cells However, as we have seen earlier, cytokinesis can also be viewed as a separate phase, which may or may not take place following mitosis. If cytokinesis does take place, cell division is not complete until the cell components have been apportioned and completely separated into the two daughter cells. Although the stages of mitosis are similar for most eukaryotes, the process of cytokinesis is quite different for eukaryotes that have cell walls, such as plant cells.

In animal cells, cytokinesis typically starts during late anaphase. A contractile ring composed of actin filaments forms just inside the plasma membrane at the former metaphase plate. The actin filaments pull

the equator of the cell inward, forming a fissure. This fissure is called the **cleavage furrow**. The furrow deepens as the actin ring contracts, and eventually the membrane is cleaved in two ([link]).

In plant cells, a new cell wall must form between the daughter cells. During interphase, the Golgi apparatus accumulates enzymes, structural proteins, and glucose molecules prior to breaking into vesicles and dispersing throughout the dividing cell. During telophase, these Golgi vesicles are transported on microtubules to form a *phragmoplast* (a vesicular structure) at the metaphase plate. There, the vesicles fuse and coalesce from the center toward the cell walls; this structure is called a **cell plate**. As more vesicles fuse, the cell plate enlarges until it merges with the cell walls at the periphery of the cell. Enzymes use the glucose that has accumulated between the membrane layers to build a new cell wall. The Golgi membranes become parts of the plasma membrane on either side of the new cell wall ([link]).



G0 Phase

Not all cells adhere to the classic cell-cycle pattern in which a newly formed daughter cell immediately enters the preparatory phases of interphase, closely followed by the mitotic phase, and cytokinesis. Cells in **G0 phase** are not actively preparing to divide. The cell is in a **quiescent** (inactive) stage that occurs when cells exit the cell cycle. Some cells enter G0 temporarily due to environmental conditions such as availability of nutrients, or

stimulation by growth factors. The cell will remain in this phase until conditions improve or until an external signal triggers the onset of G_1. Other cells that never or rarely divide, such as mature cardiac muscle and nerve cells, remain in G_0 permanently.

Scientific Method Connection

Determine the Time Spent in Cell-Cycle Stages

Problem: How long does a cell spend in interphase compared to each stage of mitosis?

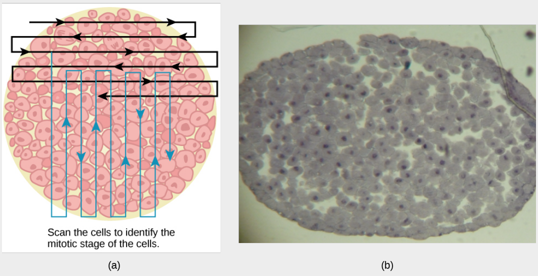
Background: A prepared microscope slide of whitefish blastula cross-sections will show cells arrested in various stages of the cell cycle. (Note: It is not visually possible to separate the stages of interphase from each other, but the mitotic stages are readily identifiable.) If 100 cells are examined, the number of cells in each identifiable cell-cycle stage will give an estimate of the time it takes for the cell to complete that stage.

Problem Statement: Given the events included in all of interphase and those that take place in each stage of mitosis, estimate the length of each stage based on a 24-hour cell cycle. Before proceeding, state your hypothesis.

Test your hypothesis: Test your hypothesis by doing the following:

1. Place a fixed and stained microscope slide of whitefish blastula cross-sections under the

scanning objective of a light microscope.
2. Locate and focus on one of the sections using the low-power objective of your microscope. Notice that the section is a circle composed of dozens of closely packed individual cells.
3. Switch to the medium-power objective and refocus. With this objective, individual cells are clearly visible, but the chromosomes will still be very small.

4. Switch to the high-power objective and slowly move the slide left to right, and up and down to view all the cells in the section ([link]). As you scan, you will notice that most of the cells are not undergoing mitosis but are in the interphase period of the cell cycle.
Slowly scan whitefish blastula cells with the high-power objective as illustrated in image (a) to identify their mitotic stage. (b) A microscopic image of the scanned cells is shown. (credit “micrograph”: modification of work by Linda Flora; scale-bar data from Matt Russell)


5. Practice identifying the various stages of the cell cycle, using the drawings of the stages as

a guide ([link]).

6. Once you are confident about your identification, begin to record the stage of each cell you encounter as you scan left to right, and top to bottom across the blastula section.
7. Keep a tally of your observations and stop when you reach 100 cells identified.
8. The larger the sample size (total number of cells counted), the more accurate the results. If possible, gather and record group data prior to calculating percentages and making estimates.

Record your observations: Make a table similar to [link] within which to record your observations.

Results of Cell Stage Identification

Phase or Stage	Individual Totals	Group Totals	Percent
Interphase			
Prophase			
Metaphase			
Anaphase			
Telophase			

Cytokinesis			
Totals	100	100	100 percent

Analyze your data/report your results: To find the length of time whitefish blastula cells spend in each stage, multiply the percent (recorded as a decimal) by 24 hours. Make a table similar to [link] to illustrate your data.

Estimate of Cell Stage Length

Phase or Stage	Percent	Time in Hours
Interphase		
Prophase		
Metaphase		
Anaphase		
Telophase		
Cytokinesis		

Draw a conclusion: Did your results support your estimated times? Were any of the outcomes unexpected? If so, discuss those events in that stage that may have contributed to the calculated time.

Section Summary

The cell cycle is an orderly sequence of events. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages. In eukaryotes, the cell cycle consists of a long preparatory period, called interphase, during which chromosomes are replicated. Interphase is divided into G_1, S, and G_2 phases. The mitotic phase begins with karyokinesis (mitosis), which consists of five stages: prophase, prometaphase, metaphase, anaphase, and telophase. The final stage of the cell division process, and sometimes viewed as the final stage of the mitotic phase, is cytokinesis, during which the cytoplasmic components of the daughter cells are separated either by an actin ring (animal cells) or by cell plate formation (plant cells).

Review Questions

Chromosomes are duplicated during what stage of the cell cycle?

1. G_1 phase
2. S phase
3. prophase

4. prometaphase

B

Which of the following events does not occur during some stages of interphase?

1. DNA duplication
2. organelle duplication
3. increase in cell size
4. separation of sister chromatids

D

The mitotic spindles arise from which cell structure?

1. centromere
2. centrosome
3. kinetochore
4. cleavage furrow

B

Attachment of the mitotic spindle fibers to the kinetochores is a characteristic of which stage

of mitosis?

1. prophase
2. prometaphase
3. metaphase
4. anaphase

B

Unpacking of chromosomes and the formation of a new nuclear envelope is a characteristic of which stage of mitosis?

1. prometaphase
2. metaphase
3. anaphase
4. telophase

D

Separation of the sister chromatids is a characteristic of which stage of mitosis?

1. prometaphase
2. metaphase
3. anaphase
4. telophase

C

The chromosomes become visible under a light microscope during which stage of mitosis?

1. prophase
2. prometaphase
3. metaphase
4. anaphase

A

The fusing of Golgi vesicles at the metaphase plate of dividing plant cells forms what structure?

1. cell plate
2. actin ring
3. cleavage furrow
4. mitotic spindle

A

[link] Which of the following is the correct order of events in mitosis?

1. Sister chromatids line up at the metaphase

plate. The kinetochore becomes attached to the mitotic spindle. The nucleus reforms and the cell divides. Cohesin proteins break down and the sister chromatids separate.

2. The kinetochore becomes attached to the mitotic spindle. Cohesin proteins break down and the sister chromatids separate. Sister chromatids line up at the metaphase plate. The nucleus reforms and the cell divides.
3. The kinetochore becomes attached to the cohesin proteins. Sister chromatids line up at the metaphase plate. The kinetochore breaks down and the sister chromatids separate. The nucleus reforms and the cell divides.
4. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

[link] D. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

Critical Thinking Questions

Briefly describe the events that occur in each phase of interphase.

During G_1, the cell increases in size, the genomic DNA is assessed for damage, and the cell stockpiles energy reserves and the components to synthesize DNA. During the S phase, the chromosomes, the centrosomes, and the centrioles (animal cells) duplicate. During the G_2 phase, the cell recovers from the S phase, continues to grow, duplicates some organelles, and dismantles other organelles.

Chemotherapy drugs such as *vincristine* (derived from Madagascar periwinkle plants) and *colchicine* (derived from autumn crocus plants) disrupt mitosis by binding to tubulin (the subunit of microtubules) and interfering with microtubule assembly and disassembly. Exactly what mitotic structure is targeted by these drugs and what effect would that have on cell division?

The mitotic spindle is formed of microtubules. Microtubules are polymers of the protein

tubulin; therefore, it is the mitotic spindle that is disrupted by these drugs. Without a functional mitotic spindle, the chromosomes will not be sorted or separated during mitosis. The cell will arrest in mitosis and die.

Describe the similarities and differences between the cytokinesis mechanisms found in animal cells versus those in plant cells.

There are very few similarities between animal cell and plant cell cytokinesis. In animal cells, a ring of actin fibers is formed around the periphery of the cell at the former metaphase plate (cleavage furrow). The actin ring contracts inward, pulling the plasma membrane toward the center of the cell until the cell is pinched in two. In plant cells, a new cell wall must be formed between the daughter cells. Due to the rigid cell walls of the parent cell, contraction of the middle of the cell is not possible. Instead, a phragmoplast first forms. Subsequently, a cell plate is formed in the center of the cell at the former metaphase plate. The cell plate is formed from Golgi vesicles that contain enzymes, proteins, and glucose. The vesicles fuse and the enzymes build a new cell wall from the proteins and glucose. The cell plate grows toward and eventually fuses with the cell wall of the parent cell.

List some reasons why a cell that has just completed cytokinesis might enter the G_0 phase instead of the G_1 phase.

Many cells temporarily enter G_0 until they reach maturity. Some cells are only triggered to enter G_1 when the organism needs to increase that particular cell type. Some cells only reproduce following an injury to the tissue. Some cells never divide once they reach maturity.

What cell-cycle events will be affected in a cell that produces mutated (non-functional) cohesin protein?

If cohesin is not functional, chromosomes are not packaged after DNA replication in the S phase of interphase. It is likely that the proteins of the centromeric region, such as the kinetochore, would not form. Even if the mitotic spindle fibers could attach to the chromatids without packing, the chromosomes would not be sorted or separated during mitosis.

Glossary

anaphase
: stage of mitosis during which sister chromatids are separated from each other

cell cycle
: ordered series of events involving cell growth and cell division that produces two new daughter cells

cell plate
: structure formed during plant cell cytokinesis by Golgi vesicles, forming a temporary structure (phragmoplast) and fusing at the metaphase plate; ultimately leads to the formation of cell walls that separate the two daughter cells

centriole
: rod-like structure constructed of microtubules at the center of each animal cell centrosome

cleavage furrow
: constriction formed by an actin ring during cytokinesis in animal cells that leads to cytoplasmic division

condensin
: proteins that help sister chromatids coil during prophase

cytokinesis
: division of the cytoplasm following mitosis that forms two daughter cells.

G_0 phase
: distinct from the G_1 phase of interphase; a cell in G_0 is not preparing to divide

G_1 phase
: (also, first gap) first phase of interphase centered on cell growth during mitosis

G_2 phase
: (also, second gap) third phase of interphase during which the cell undergoes final preparations for mitosis

interphase
: period of the cell cycle leading up to mitosis; includes G_1, S, and G_2 phases (the interim period between two consecutive cell divisions)

karyokinesis
: mitotic nuclear division

kinetochore
: protein structure associated with the centromere of each sister chromatid that attracts and binds spindle microtubules during prometaphase

metaphase plate
: equatorial plane midway between the two poles of a cell where the chromosomes align during metaphase

metaphase
: stage of mitosis during which chromosomes are aligned at the metaphase plate

mitosis
: (also, karyokinesis) period of the cell cycle during which the duplicated chromosomes are separated into identical nuclei; includes prophase, prometaphase, metaphase, anaphase, and telophase

mitotic phase
: period of the cell cycle during which duplicated chromosomes are distributed into two nuclei and cytoplasmic contents are divided; includes karyokinesis (mitosis) and cytokinesis

mitotic spindle
: apparatus composed of microtubules that orchestrates the movement of chromosomes during mitosis

prometaphase
: stage of mitosis during which the nuclear membrane breaks down and mitotic spindle fibers attach to kinetochores

prophase
: stage of mitosis during which chromosomes condense and the mitotic spindle begins to form

quiescent
: refers to a cell that is performing normal cell functions and has not initiated preparations for cell division

S phase
: second, or synthesis, stage of interphase during which DNA replication occurs

telophase
: stage of mitosis during which chromosomes arrive at opposite poles, decondense, and are surrounded by a new nuclear envelope

Control of the Cell Cycle

By the end of this section, you will be able to do the following:

- Understand how the cell cycle is controlled by mechanisms that are both internal and external to the cell
- Explain how the three internal “control checkpoints” occur at the end of G_1, at the G_2/M transition, and during metaphase
- Describe the molecules that control the cell cycle through positive and negative regulation

The length of the cell cycle is highly variable, even within the cells of a single organism. In humans, the frequency of cell turnover ranges from a few hours in early embryonic development, to an average of two to five days for epithelial cells, and to an entire human lifetime spent in G_0 by specialized cells, such as cortical neurons or cardiac muscle cells.

There is also variation in the time that a cell spends in each phase of the cell cycle. When rapidly dividing mammalian cells are grown in a culture (outside the body under optimal growing conditions), the length of the cell cycle is about 24 hours. In rapidly dividing human cells with a 24-hour cell cycle, the G_1 phase lasts approximately nine hours, the S phase lasts 10 hours, the G_2 phase lasts about four and one-half hours, and the M phase lasts approximately one-half hour. By comparison,

in fertilized eggs (and early embryos) of fruit flies, the cell cycle is completed in about eight minutes. This is because the nucleus of the fertilized egg divides many times by mitosis but does not go through cytokinesis until a multinucleate "zygote" has been produced, with many nuclei located along the periphery of the cell membrane, thereby shortening the time of the cell division cycle. The timing of events in the cell cycle of both "invertebrates" and "vertebrates" is controlled by mechanisms that are both internal and external to the cell.

Regulation of the Cell Cycle by External Events

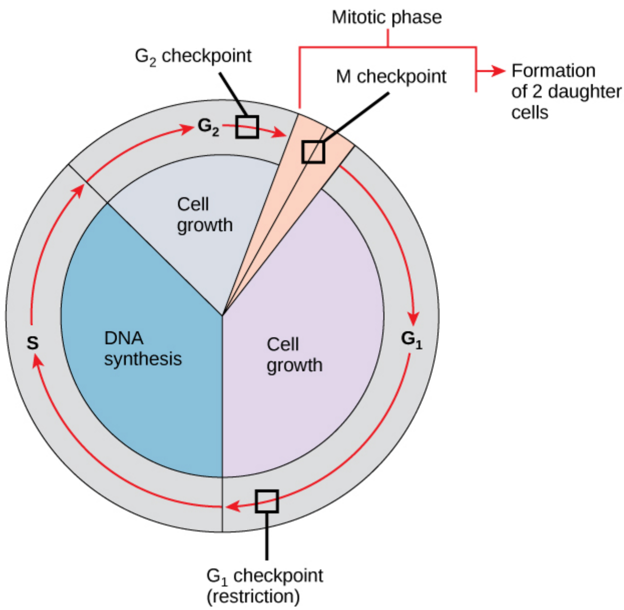
Both the initiation and inhibition of cell division are triggered by events external to the cell when it is about to begin the replication process. An event may be as simple as the death of nearby cells or as sweeping as the release of growth-promoting hormones, such as **human growth hormone (HGH or hGH)**. A lack of HGH can *inhibit* cell division, resulting in dwarfism, whereas too much HGH can result in gigantism. Crowding of cells can also inhibit cell division. In contrast, a factor that can initiate cell division is the size of the cell: As a cell grows, it becomes physiologically inefficient due to its decreasing surface-to-volume ratio. The solution to this problem is to divide.

Whatever the source of the message, the cell receives the signal, and a series of events within the cell allows it to proceed into interphase. Moving forward from this initiation point, every parameter required during each cell cycle phase must be met or the cycle cannot progress.

The cell cycle is controlled at three checkpoints. The integrity of the DNA is assessed at the G_1 checkpoint. Proper chromosome duplication is assessed at the G_2 checkpoint. Attachment of each kinetochore to a spindle fiber is assessed at the M checkpoint.

Regulation at Internal Checkpoints

It is essential that the daughter cells produced be exact duplicates of the parent cell. Mistakes in the duplication or distribution of the chromosomes lead to mutations that may be passed forward to every new cell produced from an abnormal cell. To prevent a compromised cell from continuing to divide, there are internal control mechanisms that operate at three main **cell-cycle checkpoints**: A checkpoint is one of several points in the eukaryotic cell cycle at which the progression of a cell to the next stage in the cycle can be halted until conditions are favorable. These checkpoints occur near the end of G_1, at the G_2/M transition, and during metaphase ([link]).



The G1 Checkpoint

The **G1 checkpoint** determines whether all conditions are favorable for cell division to proceed. The G1 checkpoint, also called the restriction point (in yeast), is a point at which the cell irreversibly commits to the cell division process. External influences, such as growth factors, play a large role in carrying the cell past the G1 checkpoint. In addition to adequate reserves and cell size, there is a check for genomic DNA damage at the G1 checkpoint. A cell that does not meet all the requirements will not be allowed to progress into the S phase. The cell can halt the cycle and attempt to remedy the problematic condition, or the cell can advance into G0 and await further signals when

conditions improve.

The G_2 Checkpoint

The G_2 checkpoint bars entry into the mitotic phase if certain conditions are not met. As at the G_1 checkpoint, cell size and protein reserves are assessed. However, the most important role of the G_2 checkpoint is to ensure that all of the chromosomes have been replicated and that the replicated DNA is not damaged. If the checkpoint mechanisms detect problems with the DNA, the cell cycle is halted, and the cell attempts to either complete DNA replication or repair the damaged DNA.

The M Checkpoint

The M checkpoint occurs near the end of the metaphase stage of karyokinesis. The M checkpoint is also known as the spindle checkpoint, because it determines whether all the sister chromatids are correctly attached to the spindle microtubules. Because the separation of the sister chromatids during anaphase is an irreversible step, the cycle will not proceed until the kinetochores of each pair of sister chromatids are firmly anchored to at least two spindle fibers arising from opposite poles of the cell.

Link to Learning
Watch what occurs at the G_1, G_2, and M checkpoints by visiting this website to see an animation of the cell cycle.

The concentrations of cyclin proteins change throughout the cell cycle. There is a direct correlation between cyclin accumulation and the three major cell-cycle checkpoints. Also note the sharp decline of cyclin levels following each checkpoint (the transition between phases of the cell cycle), as cyclin is degraded by cytoplasmic enzymes. (credit: modification of work by "WikiMiMa"/Wikimedia Commons)*Cyclin-dependent kinases (Cdks)* are protein kinases that, when fully activated, can phosphorylate and thus activate other proteins that advance the cell cycle past a checkpoint. To become fully activated, a Cdk must bind to a cyclin protein and then be phosphorylated by another kinase.

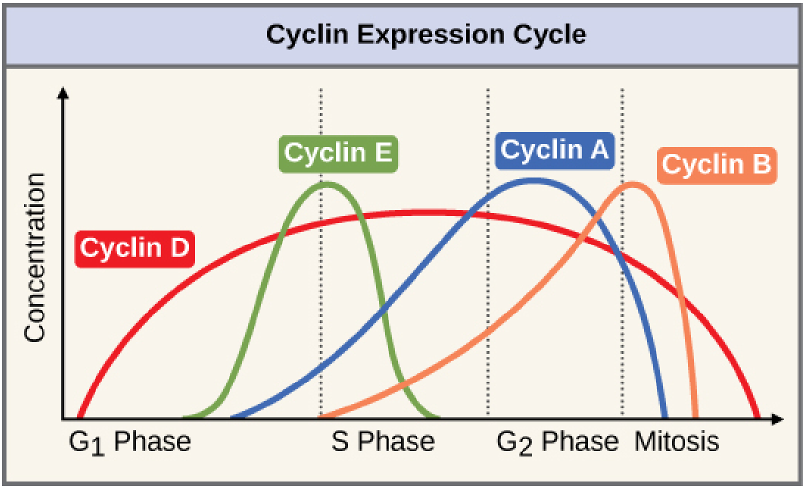
Regulator Molecules of the Cell Cycle

In addition to the internally controlled checkpoints, there are two groups of intracellular molecules that regulate the cell cycle. These regulatory molecules either promote progress of the cell to the next phase (positive regulation) or halt the cycle (negative

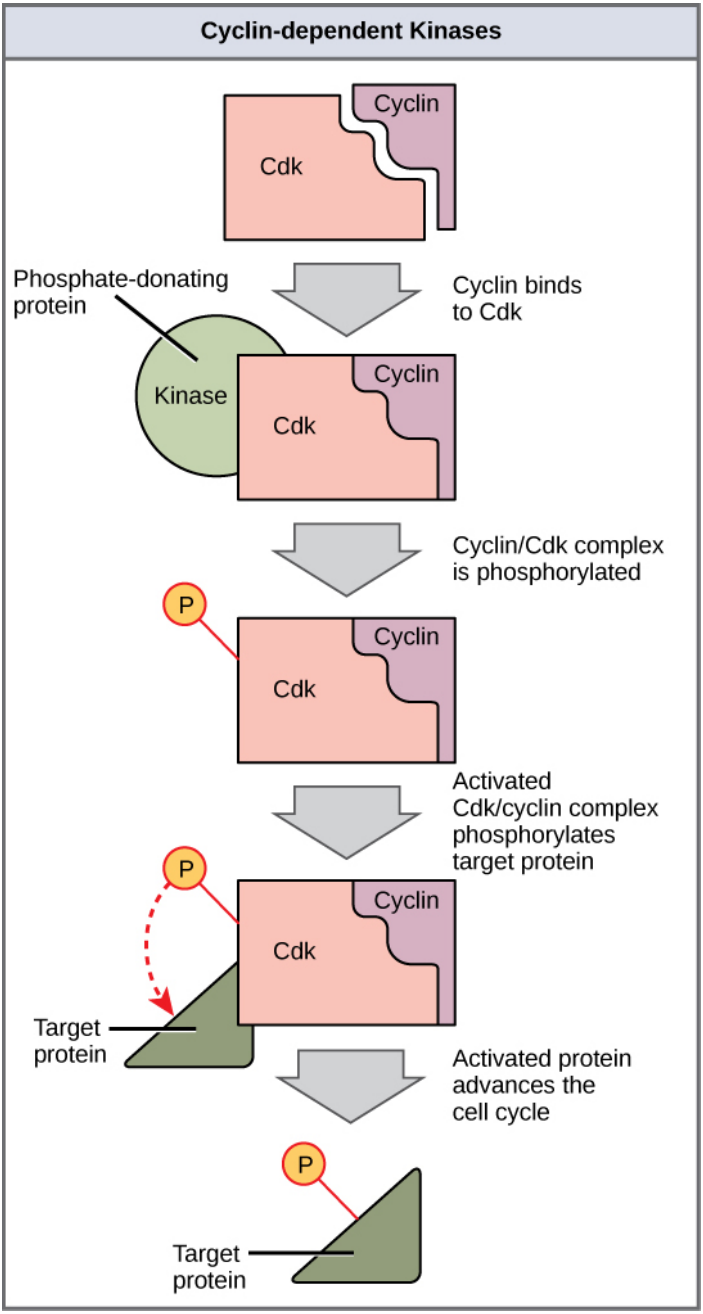
regulation). Regulator molecules may act individually, or they can influence the activity or production of other regulatory proteins. Therefore, the failure of a single regulator may have almost no effect on the cell cycle, especially if more than one mechanism controls the same event. However, the effect of a deficient or non-functioning regulator can be wide-ranging and possibly fatal to the cell if multiple processes are affected.

Positive Regulation of the Cell Cycle

Two groups of proteins, called **cyclins** and **cyclin-dependent kinases** (Cdks), are termed positive regulators. They are responsible for the progress of the cell through the various checkpoints. The levels of the four cyclin proteins fluctuate throughout the cell cycle in a predictable pattern ([link]). Increases in the concentration of cyclin proteins are triggered by both external and internal signals. After the cell moves to the next stage of the cell cycle, the cyclins that were active in the previous stage are degraded by cytoplasmic enzymes, as shown in [link] below.



Cyclins regulate the cell cycle only when they are tightly bound to Cdks. To be fully active, the Cdk/cyclin complex must also be phosphorylated in specific locations to activate the complex. Like all kinases, Cdks are enzymes (*kinases*) that in turn phosphorylate other proteins. Phosphorylation activates the protein by changing its shape. The proteins phosphorylated by Cdks are involved in advancing the cell to the next phase. ([link]). The levels of Cdk proteins are relatively stable throughout the cell cycle; however, the concentrations of cyclin fluctuate and determine when Cdk/cyclin complexes form. The different cyclins and Cdks bind at specific points in the cell cycle and thus regulate different checkpoints.



Because the cyclic fluctuations of cyclin levels are

largely based on the *timing of the cell cycle* and not on specific events, regulation of the cell cycle usually occurs by either the Cdk molecules alone or the Cdk/cyclin complexes. Without a specific concentration of fully activated cyclin/Cdk complexes, the cell cycle cannot proceed through the checkpoints.

Although the cyclins are the main regulatory molecules that determine the forward momentum of the cell cycle, there are several other mechanisms that fine-tune the progress of the cycle with negative, rather than positive, effects. These mechanisms essentially block the progression of the cell cycle until problematic conditions are resolved. Molecules that prevent the full activation of Cdks are called Cdk inhibitors. Many of these inhibitor molecules directly or indirectly monitor a particular cell-cycle event. The block placed on Cdks by inhibitor molecules will not be removed until the specific event that the inhibitor monitors is completed.

Negative Regulation of the Cell Cycle

The second group of cell-cycle regulatory molecules are *negative regulators*, which stop the cell cycle. Remember that in positive regulation, active molecules cause the cycle to progress.

The best understood negative regulatory molecules

are **retinoblastoma protein (Rb)**, **p53**, and **p21**. Retinoblastoma proteins are a group of *tumor-suppressor proteins* common in many cells. We should note here that the 53 and 21 designations refer to the functional molecular masses of the proteins (p) in kilodaltons (a dalton is equal to an *atomic mass unit*, which is equal to one proton or one neutron or 1 g/mol). Much of what is known about cell-cycle regulation comes from research conducted with cells that have *lost regulatory control*. All three of these regulatory proteins were discovered to be damaged or non-functional in cells that had begun to replicate uncontrollably (i.e., became cancerous). In each case, the main cause of the unchecked progress through the cell cycle was a faulty copy of the regulatory protein.

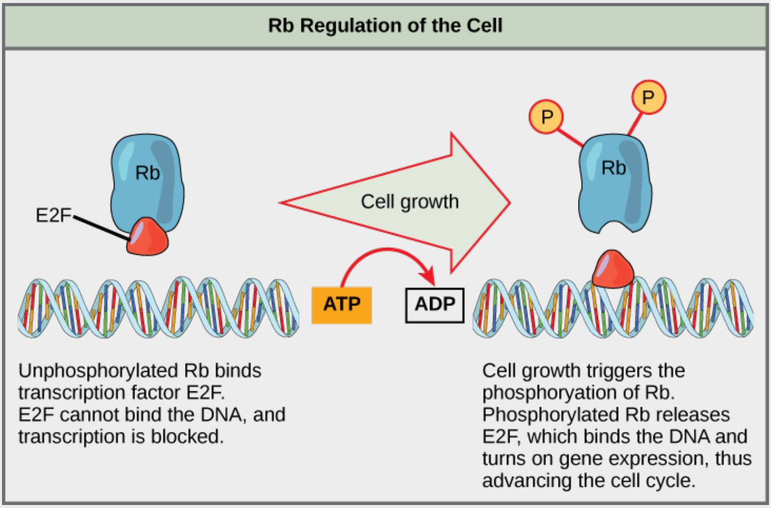
Rb, p53, and p21 act primarily at the G_1 checkpoint. p53 is a multi-functional protein that has a major impact on the commitment of a cell to division because it acts when there is damaged DNA in cells that are undergoing the preparatory processes during G_1. If damaged DNA is detected, p53 halts the cell cycle and then recruits specific enzymes to repair the DNA. If the DNA cannot be repaired, p53 can trigger apoptosis, or cell suicide, to prevent the duplication of damaged chromosomes. As p53 levels rise, the production of p21 is triggered. p21 enforces the halt in the cycle dictated by p53 by binding to and inhibiting the activity of the Cdk/cyclin complexes. As a cell is exposed to more stress,

higher levels of p53 and p21 accumulate, making it less likely that the cell will move into the S phase.

Rb, which largely monitors cell size, exerts its regulatory influence on other positive regulator proteins. In the *active*, dephosphorylated state, Rb binds to proteins called *transcription factors*, most commonly, E2F ([link]). Transcription factors “turn on” specific genes, allowing the production of proteins encoded by that gene. When Rb is bound to E2F, production of proteins necessary for the G1/S transition is blocked. As the cell increases in size, Rb is slowly phosphorylated until it becomes *inactivated*. Rb releases E2F, which can now turn on the gene that produces the transition protein, and this particular block is removed. For the cell to move past each of the checkpoints, all positive regulators must be “turned on,” and all negative regulators must be “turned off.”

Visual Connection

Rb halts the cell cycle and releases its hold in response to cell growth.



Rb and other proteins that negatively regulate the cell cycle are sometimes called tumor suppressors. Why do you think the name tumor suppressor might be appropriate for these proteins?

Section Summary

Each step of the cell cycle is monitored by internal controls called checkpoints. There are three major checkpoints in the cell cycle: one near the end of G1, a second at the G2/M transition, and the third during metaphase. Positive regulator molecules allow the cell cycle to advance to the next stage of cell division. Negative regulator molecules monitor cellular conditions and can halt the cycle until specific requirements are met.

Visual Connection Questions

[link] Rb and other proteins that negatively regulate the cell cycle are sometimes called tumor suppressors. Why do you think the name tumor suppressor might be appropriate for these proteins?

[link] Rb and other negative regulatory proteins control cell division and therefore prevent the formation of tumors. Mutations that prevent these proteins from carrying out their function can result in cancer.

Review Questions

At which of the cell-cycle checkpoints do external forces have the greatest influence?

1. G_1 checkpoint
2. G_2 checkpoint
3. M checkpoint
4. G_0 checkpoint

A

What is the main prerequisite for clearance at the G2 checkpoint?

1. cell has reached a sufficient size
2. an adequate stockpile of nucleotides
3. accurate and complete DNA replication
4. proper attachment of mitotic spindle fibers to kinetochores

---

C

If the M checkpoint is not cleared, what stage of mitosis will be blocked?

1. prophase
2. prometaphase
3. metaphase
4. anaphase

D

Which protein is a positive regulator that phosphorylates other proteins when activated?

1. p53

2. retinoblastoma protein (Rb)
3. cyclin
4. cyclin-dependent kinase (Cdk)

D

Many of the negative regulator proteins of the cell cycle were discovered in what type of cells?

1. gametes
2. cells in G_0
3. cancer cells
4. stem cells

C

Which negative regulatory molecule can trigger cell suicide (apoptosis) if vital cell cycle events do not occur?

1. p53
2. p21
3. retinoblastoma protein (Rb)
4. cyclin-dependent kinase (Cdk)

A

Critical Thinking Questions

Describe the general conditions that must be met at each of the three main cell-cycle checkpoints.

The G_1 checkpoint monitors adequate cell growth, the state of the genomic DNA, adequate stores of energy, and materials for S phase. At the G_2 checkpoint, DNA is checked to ensure that all chromosomes were duplicated and that there are no mistakes in newly synthesized DNA. Additionally, cell size and energy reserves are evaluated. The M checkpoint confirms the correct attachment of the mitotic spindle fibers to the kinetochores.

Compare and contrast the roles of the positive cell-cycle regulators negative regulators.

Positive cell regulators such as cyclin and Cdk perform tasks that advance the cell cycle to the next stage. Negative regulators such as Rb, p53, and p21 block the progression of the cell cycle until certain events have occurred.

What steps are necessary for Cdk to become fully active?

Cdk must bind to a cyclin, and it must be phosphorylated in the correct position to become fully active.

Rb is a negative regulator that blocks the cell cycle at the G_1 checkpoint until the cell achieves a requisite size. What molecular mechanism does Rb employ to halt the cell cycle?

Rb is active when it is dephosphorylated. In this state, Rb binds to E2F, which is a transcription factor required for the transcription and eventual translation of molecules required for the G_1/S transition. E2F cannot transcribe certain genes when it is bound to Rb. As the cell increases in size, Rb becomes phosphorylated, inactivated, and releases E2F. E2F can then promote the transcription of the genes it controls, and the transition proteins will be produced.

Glossary

cell-cycle checkpoint

mechanism that monitors the preparedness of a eukaryotic cell to advance through the various cell-cycle stages

cyclin
: one of a group of proteins that act in conjunction with cyclin-dependent kinases to help regulate the cell cycle by phosphorylating key proteins; the concentrations of cyclins fluctuate throughout the cell cycle

cyclin-dependent kinase (Cdk)
: one of a group of protein kinases that helps to regulate the cell cycle when bound to cyclin; it functions to phosphorylate other proteins that are either activated or inactivated by phosphorylation

p21
: cell-cycle regulatory protein that inhibits the cell cycle; its levels are controlled by p53

p53
: cell-cycle regulatory protein that regulates cell growth and monitors DNA damage; it halts the progression of the cell cycle in cases of DNA damage and may induce apoptosis

retinoblastoma protein (Rb)
: regulatory molecule that exhibits negative effects on the cell cycle by interacting with a

transcription factor (E2F)

Prokaryotic Cell Division

By the end of this section, you will be able to do the following:

- Describe the process of binary fission in prokaryotes
- Explain how FtsZ and tubulin proteins are examples of homology

Prokaryotes, such as bacteria, produce daughter cells by binary fission. For unicellular organisms, cell division is the only method to produce new individuals. In both prokaryotic and eukaryotic cells, the outcome of cell reproduction is a pair of daughter cells that are genetically identical to the parent cell. In unicellular organisms, daughter cells are individuals.

To achieve the outcome of cloned offspring, certain steps are essential. The genomic DNA must be replicated and then allocated into the daughter cells; the cytoplasmic contents must also be divided to give both new cells the cellular machinery to sustain life. As we've seen with bacterial cells, the genome consists of a single, circular DNA chromosome; therefore, the process of cell division is simplified. Karyokinesis is unnecessary because there is no true nucleus and thus no need to direct one copy of the multiple chromosomes into each daughter cell. This type of cell division is called **binary (prokaryotic) fission**.

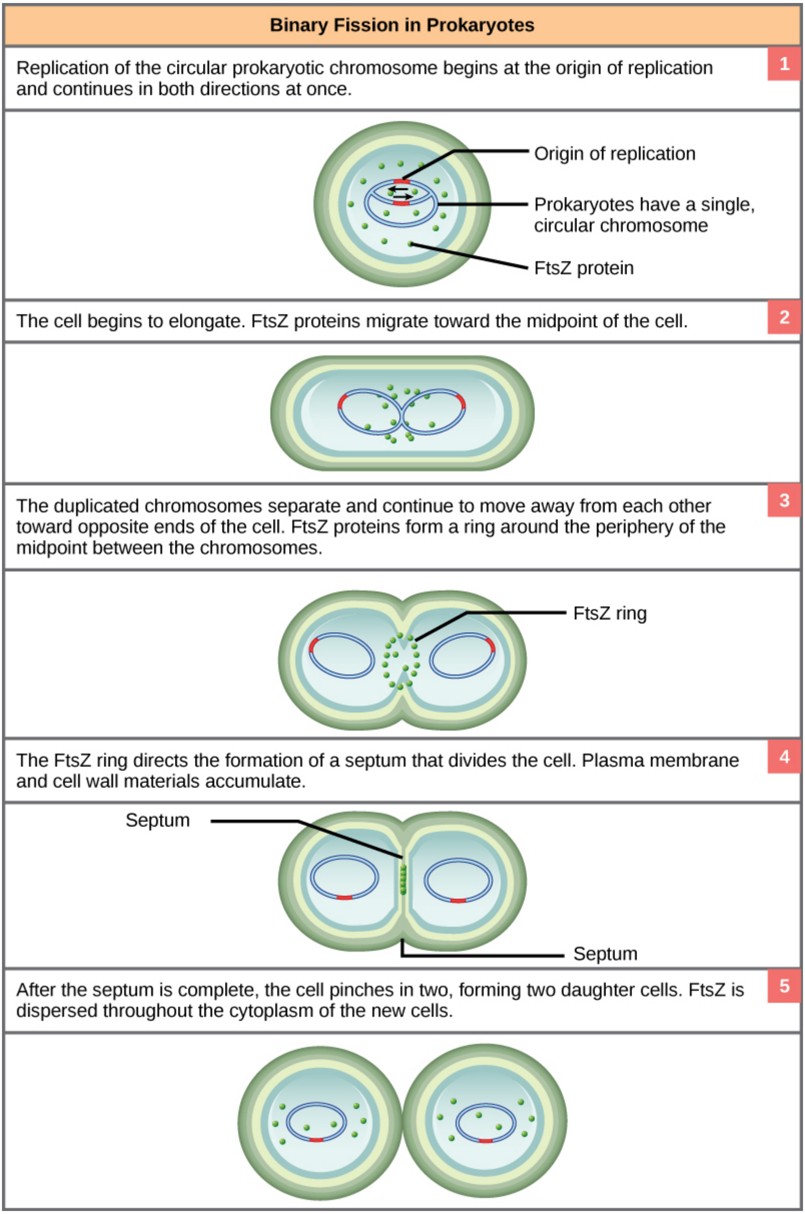
These images show the steps of binary fission in prokaryotes. (credit: modification of work by “Mcstrother”/Wikimedia Commons)

Binary Fission

Due to the relative simplicity of the prokaryotes, the cell division process is a less complicated and much more rapid process than cell division in eukaryotes. As a review of the general information on cell division we discussed at the beginning of this chapter, recall that the single, circular DNA chromosome of bacteria occupies a specific location, the nucleoid region, within the cell ([link]). Although the DNA of the nucleoid is associated with proteins that aid in packaging the molecule into a compact size, there are no histone proteins and thus no nucleosomes in prokaryotes. The packing proteins of bacteria are, however, related to the cohesin and condensin proteins involved in the chromosome compaction of eukaryotes.

The bacterial chromosome is attached to the plasma membrane at about the midpoint of the cell. The starting point of replication, the **origin**, is close to the binding site of the chromosome to the plasma membrane ([link]). Replication of the DNA is bidirectional, moving away from the origin on both strands of the loop simultaneously. As the new double strands are formed, each origin point moves away from the cell wall attachment toward the

opposite ends of the cell. As the cell elongates, the growing membrane aids in the transport of the chromosomes. After the chromosomes have cleared the midpoint of the elongated cell, cytoplasmic separation begins. The formation of a ring composed of repeating units of a protein called **FtsZ** (short for "filamenting temperature-sensitive mutant Z") directs the partition between the nucleoids. Formation of the FtsZ ring triggers the accumulation of other proteins that work together to recruit new membrane and cell wall materials to the site. A **septum** is formed between the daughter nucleoids, extending gradually from the periphery toward the center of the cell. When the new cell walls are in place, the daughter cells separate.



Evolution Connection

Mitotic Spindle Apparatus

The precise timing and formation of the mitotic spindle is critical to the success of eukaryotic cell division. Prokaryotic cells, on the other hand, do not undergo karyokinesis and therefore have no need for a mitotic spindle. However, the FtsZ protein that plays such a vital role in prokaryotic cytokinesis is structurally and functionally very similar to tubulin, the building block of the microtubules which make up the mitotic spindle fibers that are necessary for eukaryotic nuclear division. FtsZ proteins can form filaments, rings, and other three-dimensional structures that resemble the way tubulin forms microtubules, centrioles, and various cytoskeletal components. In addition, both FtsZ and tubulin employ the same energy source, GTP (guanosine triphosphate), to rapidly assemble and disassemble complex structures.

FtsZ and tubulin are considered to be homologous structures derived from common evolutionary origins. In this example, FtsZ is the ancestor protein to tubulin (an evolutionarily derived protein). While both proteins are found in extant organisms, tubulin function has evolved and diversified tremendously since evolving from its FtsZ prokaryotic origin. A survey of mitotic assembly components found in present-day unicellular eukaryotes reveals crucial intermediary steps to the complex membrane-enclosed genomes of multicellular eukaryotes ([link]).

Cell Division Apparatus among Various Organisms

	Structure of genetic material	Division of nuclear material	Separation of daughter cells
Prokaryotes	There is no nucleus. The single, circular chromosome exists in a region of cytoplasm called the nucleoid.	Occurs through binary fission. As the chromosome is replicated, the two copies move to opposite ends of the cell by an unknown mechanism.	FtsZ proteins assemble into a ring that pinches the cell in two.
Some protists	Linear chromosomes exist in the nucleus.	Chromosomes attach to the nuclear envelope, which remains intact. The mitotic	Microfilaments form a cleavage furrow that pinches the cell in two.

		spindle passes through the envelope and elongates the cell. No centrioles exist.	
Other protists	Linear chromosomes wrapped around histones exist in the nucleus.	A mitotic spindle forms from the centrioles and passes through the nuclear membrane, which remains intact. Chromosomes attach to the mitotic spindle, which separates the chromosomes and elongates	Microfilaments form a cleavage furrow that pinches the cell in two.

Animal cells	Linear chromosomes exist in the nucleus.	A mitotic spindle forms from the centrosomes. The nuclear envelope dissolves. Chromosomes attach to the mitotic spindle, which separates the chromosomes and elongates the cell.	Microfilaments form a cleavage furrow that pinches the cell in two.

Section Summary

In both prokaryotic and eukaryotic cell division, the genomic DNA is replicated and then each copy is allocated into a daughter cell. In addition, the cytoplasmic contents are divided evenly and

distributed to the new cells. However, there are many differences between prokaryotic and eukaryotic cell division. Bacteria have a single, circular DNA chromosome but no nucleus. Therefore, mitosis (karyokinesis) is not necessary in bacterial cell division. Bacterial cytokinesis is directed by a ring composed of a protein called FtsZ. Ingrowth of membrane and cell wall material from the periphery of the cells results in the formation of a septum that eventually constructs the separate cell walls of the daughter cells.

Review Questions

Which eukaryotic cell-cycle event is missing in binary fission?

1. cell growth
2. DNA duplication
3. karyokinesis
4. cytokinesis

C

FtsZ proteins direct the formation of a _____ that will eventually form the new cell walls of the daughter cells.

1. contractile ring
2. cell plate
3. cytoskeleton
4. septum

B

Critical Thinking Questions

Name the common components of eukaryotic cell division and binary fission.

The common components of eukaryotic cell division and binary fission are DNA duplication, segregation of duplicated chromosomes, and division of the cytoplasmic contents.

Describe how the duplicated bacterial chromosomes are distributed into new daughter cells without the direction of the mitotic spindle.

As the chromosome is being duplicated, each origin moves away from the starting point of replication. The chromosomes are attached to

the cell membrane via proteins; the growth of the membrane as the cell elongates aids in their movement.

Glossary

binary fission
: prokaryotic cell division process

FtsZ
: tubulin-like protein component of the prokaryotic cytoskeleton that is important in prokaryotic cytokinesis (name origin: **F**ilamenting **t**emperature-**s**ensitive mutant **Z**)

origin
: (also, ORI) region of the prokaryotic chromosome where replication begins (origin of replication)

septum
: structure formed in a bacterial cell as a precursor to the separation of the cell into two daughter cells

Introduction
class="introduction" A locust leaf consists of leaflets arrayed along a central midrib. Each leaflet is a complex photosynthetic machine, exquisitely adapted to capture sunlight and carbon dioxide. An intricate vascular system supplies the leaf with water and minerals, and exports the products of photosynthesis. (credit: modification of work by Todd Petit)

Plants are as essential to human existence as land, water, and air. Without plants, our day-to-day lives would be impossible because without oxygen from photosynthesis, aerobic life cannot be sustained. From providing food and shelter to serving as a source of medicines, oils, perfumes, and industrial products, plants provide humans with numerous valuable resources.

When you think of plants, most of the organisms that come to mind are vascular plants. These plants have tissues that conduct food and water, and they

have seeds. Seed plants are divided into gymnosperms and angiosperms. Gymnosperms include the needle-leaved conifers—spruce, fir, and pine—as well as less familiar plants, such as ginkgos and cycads. Their seeds are not enclosed by a fleshy fruit. Angiosperms, also called flowering plants, constitute the majority of seed plants. They include broadleaved trees (such as maple, oak, and elm), vegetables (such as potatoes, lettuce, and carrots), grasses, and plants known for the beauty of their flowers (roses, irises, and daffodils, for example).

While individual plant species are unique, all share a common structure: a plant body consisting of stems, roots, and leaves. They all transport water, minerals, and sugars produced through photosynthesis through the plant body in a similar manner. All plant species also respond to environmental factors, such as light, gravity, competition, temperature, and predation.

The Plant Body

By the end of this section, you will be able to do the following:

- Describe the shoot organ system and the root organ system
- Distinguish between meristematic tissue and permanent tissue
- Identify and describe the three regions where plant growth occurs
- Summarize the roles of dermal tissue, vascular tissue, and ground tissue
- Compare simple plant tissue with complex plant tissue

Like animals, plants contain cells with organelles in which specific metabolic activities take place. Unlike animals, however, plants use energy from sunlight to form sugars during photosynthesis. In addition, plant cells have cell walls, plastids, and a large central vacuole: structures that are not found in animal cells. Each of these cellular structures plays a specific role in plant structure and function.

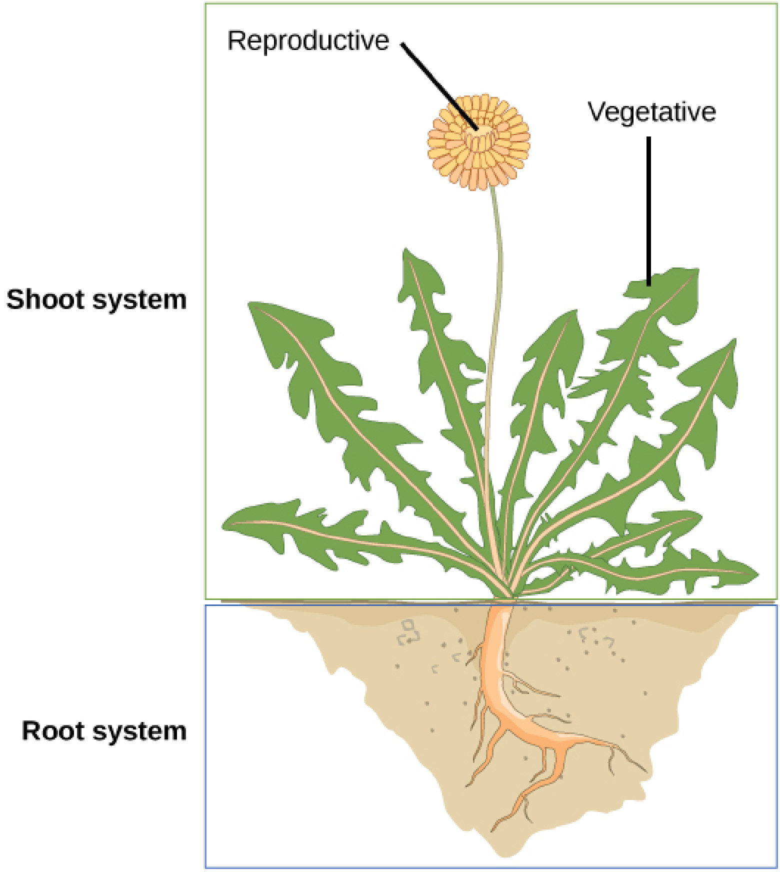
Link to Learning

Watch *Botany Without Borders*, a video produced by the Botanical Society of America about the importance of plants.

The shoot system of a plant consists of leaves, stems, flowers, and fruits. The root system anchors the plant while absorbing water and minerals from the soil.

Plant Organ Systems

In plants, just as in animals, similar cells working together form a tissue. When different types of tissues work together to perform a unique function, they form an organ; organs working together form organ systems. Vascular plants have two distinct organ systems: a shoot system, and a root system. The **shoot system** consists of two portions: the vegetative (non-reproductive) parts of the plant, such as the leaves and the stems, and the reproductive parts of the plant, which include flowers and fruits. The shoot system generally grows above ground, where it absorbs the light needed for photosynthesis. The **root system**, which supports the plants and absorbs water and minerals, is usually underground. [link] shows the organ systems of a typical plant.



This light micrograph shows a cross section of a squash (*Curcurbita maxima*) stem. Each teardrop-shaped vascular bundle consists of large xylem vessels toward the inside and smaller phloem cells toward the outside. Xylem cells, which transport water and nutrients from the roots to the rest of the plant, are dead at functional maturity. Phloem cells, which transport sugars and other organic compounds from photosynthetic tissue to the rest of the plant, are living. The vascular bundles are encased in ground tissue and surrounded by dermal tissue. (credit: modification of work by

"(biophotos)"/Flickr; scale-bar data from Matt Russell)

Plant Tissues

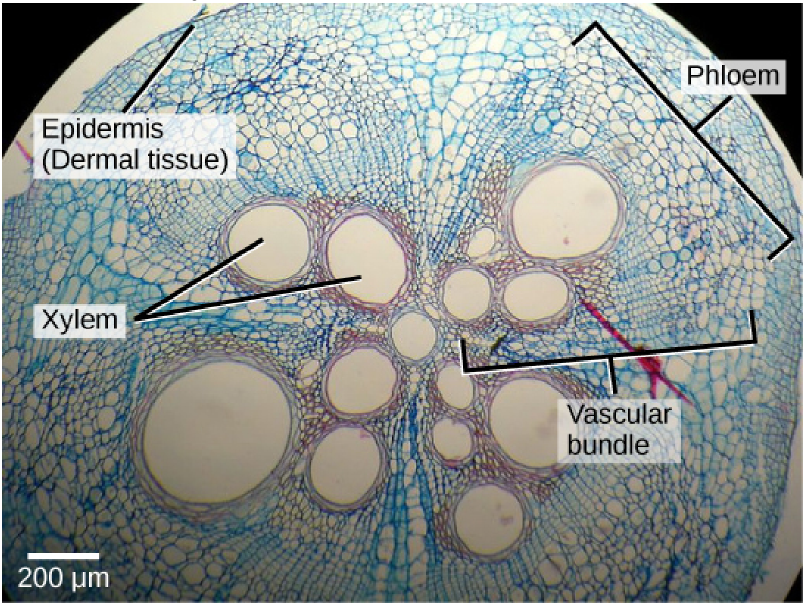
Plants are multicellular eukaryotes with tissue systems made of various cell types that carry out specific functions. Plant tissue systems fall into one of two general types: meristematic tissue, and permanent (or non-meristematic) tissue. Cells of the meristematic tissue are found in **meristems**, which are plant regions of continuous cell division and growth. **Meristematic tissue** cells are either undifferentiated or incompletely differentiated, and they continue to divide and contribute to the growth of the plant. In contrast, **permanent tissue** consists of plant cells that are no longer actively dividing.

Meristematic tissues consist of three types, based on their location in the plant. **Apical meristems** contain meristematic tissue located at the tips of stems and roots, which enable a plant to extend in length. **Lateral meristems** facilitate growth in thickness or girth in a maturing plant. **Intercalary meristems** occur only in monocots, at the bases of leaf blades and at nodes (the areas where leaves attach to a stem). This tissue enables the monocot leaf blade to increase in length from the leaf base; for example, it allows lawn grass leaves to elongate even after repeated mowing.

Meristems produce cells that quickly differentiate, or specialize, and become permanent tissue. Such cells take on specific roles and lose their ability to divide further. They differentiate into three main types: dermal, vascular, and ground tissue. **Dermal tissue** covers and protects the plant, and **vascular tissue** transports water, minerals, and sugars to different parts of the plant. **Ground tissue** serves as a site for photosynthesis, provides a supporting matrix for the vascular tissue, and helps to store water and sugars.

Secondary tissues are either simple (composed of similar cell types) or complex (composed of different cell types). Dermal tissue, for example, is a simple tissue that covers the outer surface of the plant and controls gas exchange. Vascular tissue is an example of a complex tissue, and is made of two specialized conducting tissues: xylem and phloem. Xylem tissue transports water and nutrients from the roots to different parts of the plant, and includes three different cell types: vessel elements and tracheids (both of which conduct water), and xylem parenchyma. Phloem tissue, which transports organic compounds from the site of photosynthesis to other parts of the plant, consists of four different cell types: sieve cells (which conduct photosynthates), companion cells, phloem parenchyma, and phloem fibers. Unlike xylem conducting cells, phloem conducting cells are alive at maturity. The xylem and phloem always lie

adjacent to each other ([link]). In stems, the xylem and the phloem form a structure called a **vascular bundle**; in roots, this is termed the **vascular stele** or **vascular cylinder**.



Section Summary

A vascular plant consists of two organ systems: the shoot system and the root system. The shoot system includes the aboveground vegetative portions (stems and leaves) and reproductive parts (flowers and fruits). The root system supports the plant and is usually underground. A plant is composed of two main types of tissue: meristematic tissue and permanent tissue. Meristematic tissue consists of actively dividing cells found in root and shoot tips.

As growth occurs, meristematic tissue differentiates into permanent tissue, which is categorized as either simple or complex. Simple tissues are made up of similar cell types; examples include dermal tissue and ground tissue. Dermal tissue provides the outer covering of the plant. Ground tissue is responsible for photosynthesis; it also supports vascular tissue and may store water and sugars. Complex tissues are made up of different cell types. Vascular tissue, for example, is made up of xylem and phloem cells.

Review Questions

Plant regions of continuous growth are made up of ______.

1. dermal tissue
2. vascular tissue
3. meristematic tissue
4. permanent tissue

C

Which of the following is the major site of photosynthesis?

1. apical meristem

2. ground tissue
3. xylem cells
4. phloem cells

B

Critical Thinking Questions

What type of meristem is found only in monocots, such as lawn grasses? Explain how this type of meristematic tissue is beneficial in lawn grasses that are mowed each week.

Lawn grasses and other monocots have an intercalary meristem, which is a region of meristematic tissue at the base of the leaf blade. This is beneficial to the plant because it can continue to grow even when the tip of the plant is removed by grazing or mowing.

Which plant part is responsible for transporting water, minerals, and sugars to different parts of the plant? Name the two types of tissue that make up this overall tissue, and explain the role of each.

Vascular tissue transports water, minerals, and sugars throughout the plant. Vascular tissue is made up of xylem tissue and phloem tissue. Xylem tissue transports water and nutrients from the roots upward. Phloem tissue carries sugars from the sites of photosynthesis to the rest of the plant.

Glossary

apical meristem
: meristematic tissue located at the tips of stems and roots; enables a plant to extend in length

dermal tissue
: protective plant tissue covering the outermost part of the plant; controls gas exchange

ground tissue
: plant tissue involved in photosynthesis; provides support, and stores water and sugars

intercalary meristem
: meristematic tissue located at nodes and the bases of leaf blades; found only in monocots

lateral meristem
: meristematic tissue that enables a plant to increase in thickness or girth

meristematic tissue
: tissue containing cells that constantly divide; contributes to plant growth

meristem
: plant region of continuous growth

permanent tissue
: plant tissue composed of cells that are no longer actively dividing

root system
: belowground portion of the plant that supports the plant and absorbs water and minerals

shoot system
: aboveground portion of the plant; consists of nonreproductive plant parts, such as leaves and stems, and reproductive parts, such as flowers and fruits

vascular bundle
: strands of stem tissue made up of xylem and phloem

vascular stele
: strands of root tissue made up of xylem and phloem

vascular tissue
: tissue made up of xylem and phloem that

transports food and water throughout the plant

Stems

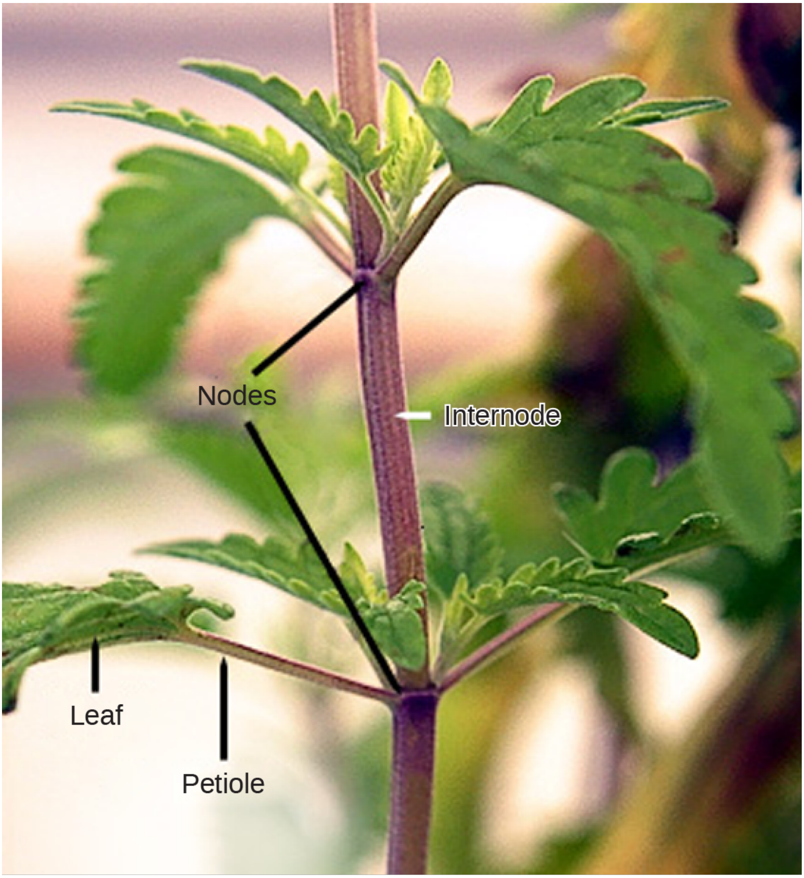
By the end of this section, you will be able to do the following:

- Describe the main function and basic structure of stems
- Compare and contrast the roles of dermal tissue, vascular tissue, and ground tissue
- Distinguish between primary growth and secondary growth in stems
- Summarize the origin of annual rings
- List and describe examples of modified stems

Stems are a part of the shoot system of a plant. They may range in length from a few millimeters to hundreds of meters, and also vary in diameter, depending on the plant type. Stems are usually above ground, although the stems of some plants, such as the potato, also grow underground. Stems may be herbaceous (soft) or woody in nature. Their main function is to provide support to the plant, holding leaves, flowers and buds; in some cases, stems also store food for the plant. A stem may be unbranched, like that of a palm tree, or it may be highly branched, like that of a magnolia tree. The stem of the plant connects the roots to the leaves, helping to transport absorbed water and minerals to different parts of the plant. It also helps to transport the products of photosynthesis, namely sugars, from the leaves to the rest of the plant.

Plant stems, whether above or below ground, are characterized by the presence of nodes and internodes ([link]). **Nodes** are points of attachment for leaves, aerial roots, and flowers. The stem region between two nodes is called an **internode**. The stalk that extends from the stem to the base of the leaf is the petiole. An **axillary bud** is usually found in the axil—the area between the base of a leaf and the stem—where it can give rise to a branch or a flower. The apex (tip) of the shoot contains the apical meristem within the **apical bud**.

Leaves are attached to the plant stem at areas called nodes. An internode is the stem region between two nodes. The petiole is the stalk connecting the leaf to the stem. The leaves just above the nodes arose from axillary buds.



The stem of common St John's Wort (*Hypericum perforatum*) is shown in cross section in this light micrograph. The central pith (greenish-blue, in the center) and peripheral cortex (narrow zone 3–5 cells thick just inside the epidermis) are composed of parenchyma cells. Vascular tissue composed of xylem (red) and phloem tissue (green, between the xylem and cortex) surrounds the pith. (credit: Rolf-Dieter Mueller) Collenchyma cell walls are uneven in thickness, as seen in this light micrograph. They provide support to plant structures. (credit: modification of work by Carl Szczerski; scale-bar

data from Matt Russell)Openings called stomata (singular: stoma) allow a plant to take up carbon dioxide and release oxygen and water vapor. The (a) colorized scanning-electron micrograph shows a closed stoma of a dicot. Each stoma is flanked by two guard cells that regulate its (b) opening and closing. The (c) guard cells sit within the layer of epidermal cells. (credit a: modification of work by Louisa Howard, Rippel Electron Microscope Facility, Dartmouth College; credit b: modification of work by June Kwak, University of Maryland; scale-bar data from Matt Russell) In (a) dicot stems, vascular bundles are arranged around the periphery of the ground tissue. The xylem tissue is located toward the interior of the vascular bundle, and phloem is located toward the exterior. Sclerenchyma fibers cap the vascular bundles. In (b) monocot stems, vascular bundles composed of xylem and phloem tissues are scattered throughout the ground tissue.

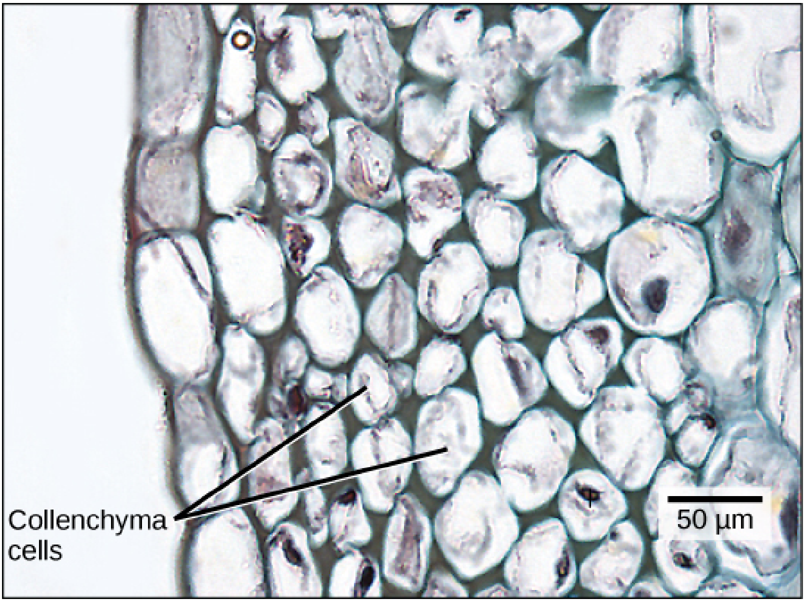
Stem Anatomy

The stem and other plant organs arise from the ground tissue, and are primarily made up of simple tissues formed from three types of cells: parenchyma, collenchyma, and sclerenchyma cells.

Parenchyma cells are the most common plant cells ([link]). They are found in the stem, the root, the inside of the leaf, and the pulp of the fruit. Parenchyma cells are responsible for metabolic

functions, such as photosynthesis, and they help repair and heal wounds. Some parenchyma cells also store starch.

Collenchyma cells are elongated cells with unevenly thickened walls ([link]). They provide structural support, mainly to the stem and leaves. These cells are alive at maturity and are usually found below the epidermis. The “strings” of a celery stalk are an example of collenchyma cells.

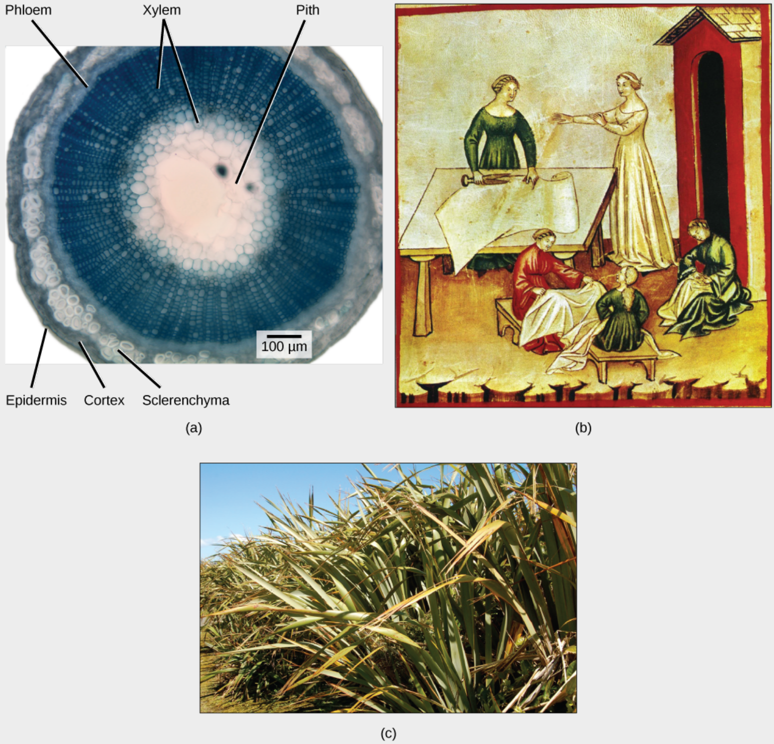


Sclerenchyma cells also provide support to the plant, but unlike collenchyma cells, many of them are dead at maturity. There are two types of sclerenchyma cells: fibers and sclereids. Both types have secondary cell walls that are thickened with deposits of lignin, an organic compound that is a key component of wood. Fibers are long, slender cells; sclereids are smaller-sized. Sclereids give pears their gritty texture. Humans use sclerenchyma fibers to make linen and rope ([link]).

Visual Connection

The central pith and outer cortex of the (a) flax stem are made up of parenchyma cells. Inside the cortex is a layer of sclerenchyma cells, which make

up the fibers in flax rope and clothing. Humans have grown and harvested flax for thousands of years. In (b) this drawing, fourteenth-century women prepare linen. The (c) flax plant is grown and harvested for its fibers, which are used to weave linen, and for its seeds, which are the source of linseed oil. (credit a: modification of work by Emmanuel Boutet based on original work by Ryan R. MacKenzie; credit c: modification of work by Brian Dearth; scale-bar data from Matt Russell)



Which layers of the stem are made of parenchyma cells?

1. cortex and pith
2. phloem

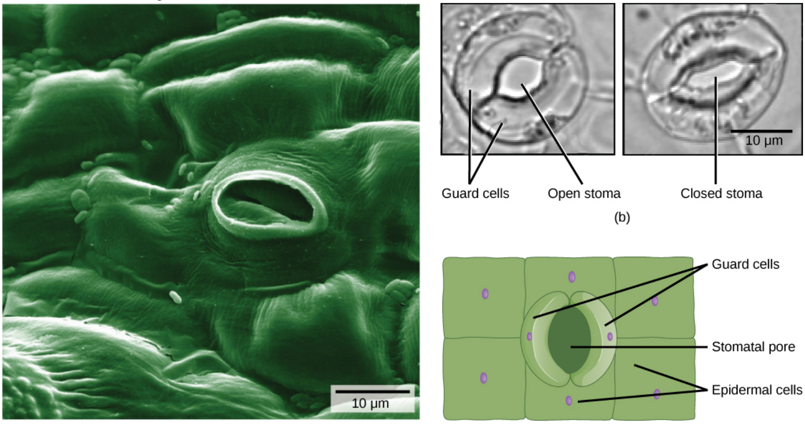
3. sclerenchyma
4. xylem

Like the rest of the plant, the stem has three tissue systems: dermal, vascular, and ground tissue. Each is distinguished by characteristic cell types that perform specific tasks necessary for the plant's growth and survival.

Dermal Tissue

The dermal tissue of the stem consists primarily of **epidermis**, a single layer of cells covering and protecting the underlying tissue. Woody plants have a tough, waterproof outer layer of cork cells commonly known as **bark**, which further protects the plant from damage. Epidermal cells are the most numerous and least differentiated of the cells in the epidermis. The epidermis of a leaf also contains openings known as stomata, through which the exchange of gases takes place ([link]). Two cells, known as **guard cells**, surround each leaf stoma, controlling its opening and closing and thus regulating the uptake of carbon dioxide and the release of oxygen and water vapor. **Trichomes** are hair-like structures on the epidermal surface. They help to reduce **transpiration** (the loss of water by aboveground plant parts), increase solar reflectance, and store compounds that defend the leaves against

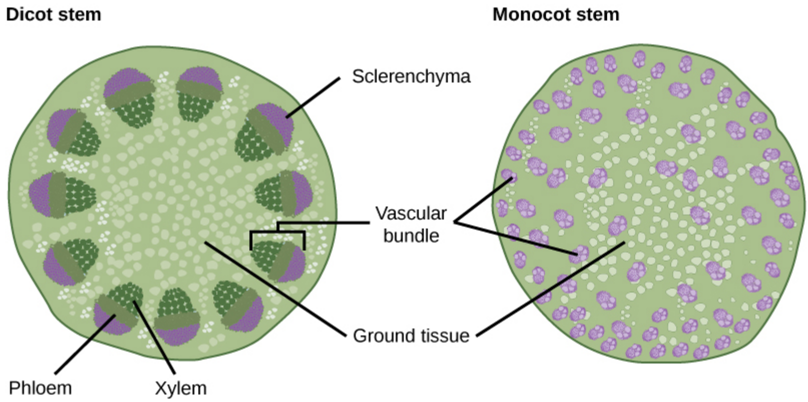
predation by herbivores.



(a) (b) (c)

Vascular Tissue

The xylem and phloem that make up the vascular tissue of the stem are arranged in distinct strands called vascular bundles, which run up and down the length of the stem. When the stem is viewed in cross section, the vascular bundles of dicot stems are arranged in a ring. In plants with stems that live for more than one year, the individual bundles grow together and produce the characteristic growth rings. In monocot stems, the vascular bundles are randomly scattered throughout the ground tissue ([link]).



Xylem tissue has three types of cells: xylem parenchyma, tracheids, and vessel elements. The latter two types conduct water and are dead at maturity. **Tracheids** are xylem cells with thick secondary cell walls that are lignified. Water moves from one tracheid to another through regions on the side walls known as pits, where secondary walls are absent. **Vessel elements** are xylem cells with thinner walls; they are shorter than tracheids. Each vessel element is connected to the next by means of a perforation plate at the end walls of the element. Water moves through the perforation plates to travel up the plant.

Phloem tissue is composed of sieve-tube cells, companion cells, phloem parenchyma, and phloem fibers. A series of **sieve-tube cells** (also called sieve-tube elements) are arranged end to end to make up a long sieve tube, which transports organic substances such as sugars and amino acids. The sugars flow from one sieve-tube cell to the next through perforated sieve plates, which are found at

the end junctions between two cells. Although still alive at maturity, the nucleus and other cell components of the sieve-tube cells have disintegrated. **Companion cells** are found alongside the sieve-tube cells, providing them with metabolic support. The companion cells contain more ribosomes and mitochondria than the sieve-tube cells, which lack some cellular organelles.

Ground Tissue

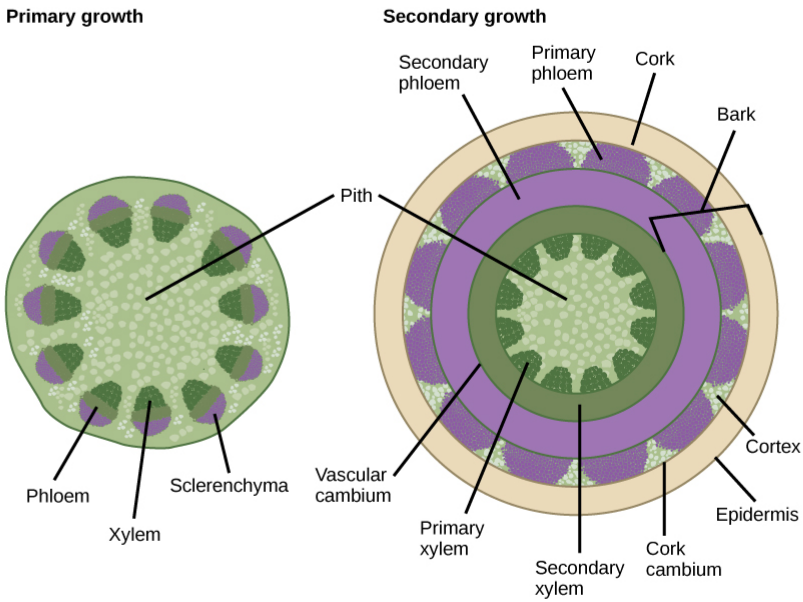
Ground tissue is mostly made up of parenchyma cells, but may also contain collenchyma and sclerenchyma cells that help support the stem. The ground tissue towards the interior of the vascular tissue in a stem or root is known as **pith**, while the layer of tissue between the vascular tissue and the epidermis is known as the **cortex**.

In woody plants, primary growth is followed by secondary growth, which allows the plant stem to increase in thickness or girth. Secondary vascular tissue is added as the plant grows, as well as a cork layer. The bark of a tree extends from the vascular cambium to the epidermis. Lenticels on the bark of this cherry tree enable the woody stem to exchange gases with the surrounding atmosphere. (credit: Roger Griffith) The rate of wood growth increases in summer and decreases in winter, producing a characteristic ring for each year of growth. Seasonal changes in weather patterns can also affect the growth rate—note how the rings vary in thickness.

(credit: Adrian Pingstone)

Growth in Stems

Growth in plants occurs as the stems and roots lengthen. Some plants, especially those that are woody, also increase in thickness during their life span. The increase in length of the shoot and the root is referred to as **primary growth**, and is the result of cell division in the shoot apical meristem. **Secondary growth** is characterized by an increase in thickness or girth of the plant, and is caused by cell division in the lateral meristem. [link] shows the areas of primary and secondary growth in a plant. Herbaceous plants mostly undergo primary growth, with hardly any secondary growth or increase in thickness. Secondary growth or "wood" is noticeable in woody plants; it occurs in some dicots, but occurs very rarely in monocots.



Some plant parts, such as stems and roots, continue to grow throughout a plant's life: a phenomenon called indeterminate growth. Other plant parts, such as leaves and flowers, exhibit determinate growth, which ceases when a plant part reaches a particular size.

Primary Growth

Most primary growth occurs at the apices, or tips, of stems and roots. Primary growth is a result of rapidly dividing cells in the apical meristems at the shoot tip and root tip. Subsequent cell elongation also contributes to primary growth. The growth of shoots and roots during primary growth enables plants to continuously seek water (roots) or sunlight (shoots).

The influence of the apical bud on overall plant growth is known as apical dominance, which diminishes the growth of axillary buds that form along the sides of branches and stems. Most coniferous trees exhibit strong apical dominance, thus producing the typical conical Christmas tree shape. If the apical bud is removed, then the axillary buds will start forming lateral branches. Gardeners make use of this fact when they prune plants by cutting off the tops of branches, thus encouraging the axillary buds to grow out, giving the plant a bushy shape.

> Link to Learning
> Watch this BBC Nature video showing how time-lapse photography captures plant growth at high speed.

Secondary Growth

The increase in stem thickness that results from secondary growth is due to the activity of the lateral meristems, which are lacking in herbaceous plants. Lateral meristems include the vascular cambium and, in woody plants, the cork cambium (see [link]). The vascular cambium is located just outside the primary xylem and to the interior of the primary

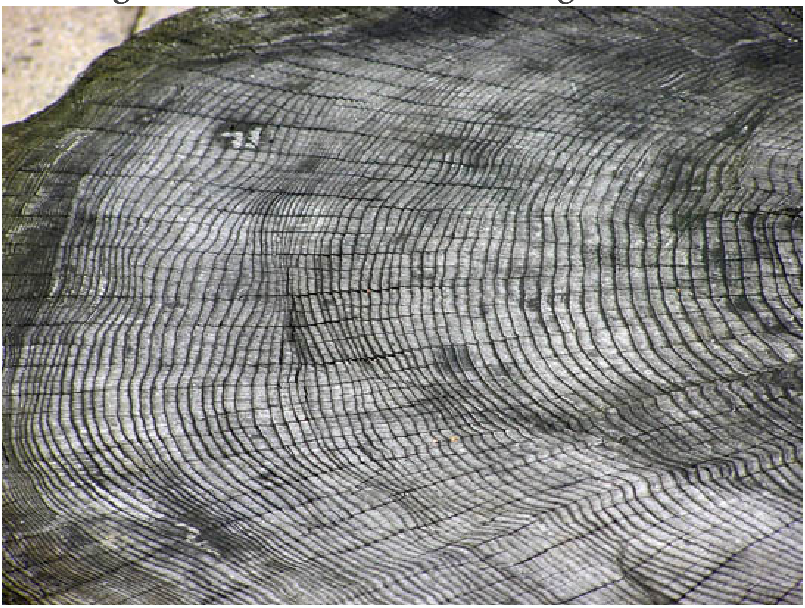
phloem. The cells of the vascular cambium divide and form secondary xylem (tracheids and vessel elements) to the inside, and secondary phloem (sieve elements and companion cells) to the outside. The thickening of the stem that occurs in secondary growth is due to the formation of secondary phloem and secondary xylem by the vascular cambium, plus the action of cork cambium, which forms the tough outermost layer of the stem. The cells of the secondary xylem contain lignin, which provides hardiness and strength.

In woody plants, cork cambium is the outermost lateral meristem. It produces cork cells (bark) containing a waxy substance known as suberin that can repel water. The bark protects the plant against physical damage and helps reduce water loss. The cork cambium also produces a layer of cells known as phelloderm, which grows inward from the cambium. The cork cambium, cork cells, and phelloderm are collectively termed the **periderm**. The periderm substitutes for the epidermis in mature plants. In some plants, the periderm has many openings, known as **lenticels**, which allow the interior cells to exchange gases with the outside atmosphere ([link]). This supplies oxygen to the living and metabolically active cells of the cortex, xylem, and phloem.

Annual Rings

The activity of the vascular cambium gives rise to annual growth rings. During the spring growing season, cells of the secondary xylem have a large internal diameter and their primary cell walls are not extensively thickened. This is known as early wood, or spring wood. During the fall season, the secondary xylem develops thickened cell walls, forming late wood, or autumn wood, which is

denser than early wood. This alternation of early and late wood is due largely to a seasonal decrease in the number of vessel elements and a seasonal increase in the number of tracheids. It results in the formation of an annual ring, which can be seen as a circular ring in the cross section of the stem ([link]). An examination of the number of annual rings and their nature (such as their size and cell wall thickness) can reveal the age of the tree and the prevailing climatic conditions during each season.



Stem modifications enable plants to thrive in a variety of environments. Shown are (a) ginger (*Zingiber officinale*) rhizomes, (b) a carrion flower (*Amorphophallus titanum*) corm, (c) Rhodes grass (*Chloris gayana*) stolons, (d) strawberry (*Fragaria ananassa*) runners, (e) potato (*Solanum tuberosum*) tubers, and (f) red onion (*Allium*) bulbs. (credit a: modification of work by Maja Dumat; credit c:

modification of work by Harry Rose; credit d: modification of work by Rebecca Siegel; credit e: modification of work by Scott Bauer, USDA ARS; credit f: modification of work by Stephen Ausmus, USDA ARS) Found in southeastern United States, (a) buckwheat vine (*Brunnichia ovata*) is a weedy plant that climbs with the aid of tendrils. This one is shown climbing up a wooden stake. (b) Thorns are modified branches. (credit a: modification of work by Christopher Meloche, USDA ARS; credit b: modification of work by "macrophile"/Flickr)

Stem Modifications

Some plant species have modified stems that are especially suited to a particular habitat and environment ([link]). A **rhizome** is a modified stem that grows horizontally underground and has nodes and internodes. Vertical shoots may arise from the buds on the rhizome of some plants, such as ginger and ferns. **Corms** are similar to rhizomes, except they are more rounded and fleshy (such as in gladiolus). Corms contain stored food that enables some plants to survive the winter. **Stolons** are stems that run almost parallel to the ground, or just below the surface, and can give rise to new plants at the nodes. **Runners** are a type of stolon that runs above the ground and produces new clone plants at nodes at varying intervals: strawberries are an example. **Tubers** are modified stems that may store starch, as seen in the potato (*Solanum* sp.). Tubers arise as

swollen ends of stolons, and contain many adventitious or unusual buds (familiar to us as the "eyes" on potatoes). A **bulb**, which functions as an underground storage unit, is a modification of a stem that has the appearance of enlarged fleshy leaves emerging from the stem or surrounding the base of the stem, as seen in the iris.

Link to Learning

Watch botanist Wendy Hodgson, of Desert Botanical Garden in Phoenix, Arizona, explain how agave plants were cultivated for food hundreds of years ago in the Arizona desert in this video: *Finding the Roots of an Ancient Crop*.

Some aerial modifications of stems are tendrils and thorns ([link]). **Tendrils** are slender, twining strands that enable a plant (like a vine or pumpkin) to seek support by climbing on other surfaces.

Thorns are modified branches appearing as sharp outgrowths that protect the plant; common examples include roses, Osage orange, and devil's walking stick.

(a) (b)

Section Summary

The stem of a plant bears the leaves, flowers, and fruits. Stems are characterized by the presence of nodes (the points of attachment for leaves or branches) and internodes (regions between nodes).

Plant organs are made up of simple and complex tissues. The stem has three tissue systems: dermal, vascular, and ground tissue. Dermal tissue is the outer covering of the plant. It contains epidermal cells, stomata, guard cells, and trichomes. Vascular tissue is made up of xylem and phloem tissues and conducts water, minerals, and photosynthetic products. Ground tissue is responsible for

photosynthesis and support and is composed of parenchyma, collenchyma, and sclerenchyma cells.

Primary growth occurs at the tips of roots and shoots, causing an increase in length. Woody plants may also exhibit secondary growth, or increase in thickness. In woody plants, especially trees, annual rings may form as growth slows at the end of each season. Some plant species have modified stems that help to store food, propagate new plants, or discourage predators. Rhizomes, corms, stolons, runners, tubers, bulbs, tendrils, and thorns are examples of modified stems.

Visual Connection Questions

[link] Which layers of the stem are made of parenchyma cells?

1. cortex and pith
2. epidermis
3. sclerenchyma
4. epidermis and cortex

[link] A and B. The cortex, pith, and epidermis are made of parenchyma cells.

Review Questions

Stem regions at which leaves are attached are called _____.

1. trichomes
2. lenticels
3. nodes
4. internodes

C

Which of the following cell types forms most of the inside of a plant?

1. meristem cells
2. collenchyma cells
3. sclerenchyma cells
4. parenchyma cells

D

Tracheids, vessel elements, sieve-tube cells, and companion cells are components of _____.

1. vascular tissue
2. meristematic tissue

3. ground tissue
4. dermal tissue

A

The primary growth of a plant is due to the action of the ______.

1. lateral meristem
2. vascular cambium
3. apical meristem
4. cork cambium

C

Which of the following is an example of secondary growth?

1. increase in length
2. increase in thickness or girth
3. increase in root hairs
4. increase in leaf number

B

Secondary growth in stems is usually seen in

_____.

1. monocots
2. dicots
3. both monocots and dicots
4. neither monocots nor dicots

B

Critical Thinking Questions

Describe the roles played by stomata and guard cells. What would happen to a plant if these cells did not function correctly?

Stomata allow gases to enter and exit the plant. Guard cells regulate the opening and closing of stomata. If these cells did not function correctly, a plant could not get the carbon dioxide needed for photosynthesis, nor could it release the oxygen produced by photosynthesis.

Compare the structure and function of xylem to that of phloem.

Xylem is made up tracheids and vessel elements, which are cells that transport water and dissolved minerals and that are dead at maturity. Phloem is made up of sieve-tube cells and companion cells, which transport carbohydrates and are alive at maturity.

Explain the role of the cork cambium in woody plants.

In woody plants, the cork cambium is the outermost lateral meristem; it produces new cells towards the interior, which enables the plant to increase in girth. The cork cambium also produces cork cells towards the exterior, which protect the plant from physical damage while reducing water loss.

What is the function of lenticels?

In woody stems, lenticels allow internal cells to exchange gases with the outside atmosphere.

Besides the age of a tree, what additional information can annual rings reveal?

Annual rings can also indicate the climate conditions that prevailed during each growing season.

Give two examples of modified stems and explain how each example benefits the plant.

Answers will vary. Rhizomes, stolons, and runners can give rise to new plants. Corms, tubers, and bulbs can also produce new plants and can store food. Tendrils help a plant to climb, while thorns discourage herbivores.

Glossary

apical bud
: bud formed at the tip of the shoot

axillary bud
: bud located in the axil: the stem area where the petiole connects to the stem

bark
: tough, waterproof, outer epidermal layer of cork cells

bulb
: modified underground stem that consists of a large bud surrounded by numerous leaf scales

collenchyma cell
: elongated plant cell with unevenly thickened walls; provides structural support to the stem and leaves

companion cell
: phloem cell that is connected to sieve-tube cells; has large amounts of ribosomes and mitochondria

corm
: rounded, fleshy underground stem that contains stored food

cortex
: ground tissue found between the vascular tissue and the epidermis in a stem or root

epidermis
: single layer of cells found in plant dermal tissue; covers and protects underlying tissue

guard cells
: paired cells on either side of a stoma that control stomatal opening and thereby regulate the movement of gases and water vapor

internode
: region between nodes on the stem

lenticel

opening on the surface of mature woody stems that facilitates gas exchange

node
: point along the stem at which leaves, flowers, or aerial roots originate

parenchyma cell
: most common type of plant cell; found in the stem, root, leaf, and in fruit pulp; site of photosynthesis and starch storage

periderm
: outermost covering of woody stems; consists of the cork cambium, cork cells, and the phelloderm

pith
: ground tissue found towards the interior of the vascular tissue in a stem or root

primary growth
: growth resulting in an increase in length of the stem and the root; caused by cell division in the shoot or root apical meristem

rhizome
: modified underground stem that grows horizontally to the soil surface and has nodes and internodes

runner

stolon that runs above the ground and produces new clone plants at nodes

sclerenchyma cell
: plant cell that has thick secondary walls and provides structural support; usually dead at maturity

secondary growth
: growth resulting in an increase in thickness or girth; caused by the lateral meristem and cork cambium

sieve-tube cell
: phloem cell arranged end to end to form a sieve tube that transports organic substances such as sugars and amino acids

stolon
: modified stem that runs parallel to the ground and can give rise to new plants at the nodes

tendril
: modified stem consisting of slender, twining strands used for support or climbing

thorn
: modified stem branch appearing as a sharp outgrowth that protects the plant

tracheid
: xylem cell with thick secondary walls that

helps transport water

trichome
: hair-like structure on the epidermal surface

tuber
: modified underground stem adapted for starch storage; has many adventitious buds

vessel element
: xylem cell that is shorter than a tracheid and has thinner walls

Roots
By the end of this section, you will be able to do the following:

- Identify the two types of root systems
- Describe the three zones of the root tip and summarize the role of each zone in root growth
- Describe the structure of the root
- List and describe examples of modified roots

The roots of seed plants have three major functions: anchoring the plant to the soil, absorbing water and minerals and transporting them upwards, and storing the products of photosynthesis. Some roots are modified to absorb moisture and exchange gases. Most roots are underground. Some plants, however, also have **adventitious roots**, which emerge above the ground from the shoot.
(a) Tap root systems have a main root that grows down, while (b) fibrous root systems consist of many small roots. (credit b: modification of work by "Austen Squarepants"/Flickr)

Types of Root Systems

Root systems are mainly of two types ([link]). Dicots have a tap root system, while monocots have a fibrous root system. A **tap root system** has a main root that grows down vertically, and from which many smaller lateral roots arise. Dandelions are a

good example; their tap roots usually break off when trying to pull these weeds, and they can regrow another shoot from the remaining root. A tap root system penetrates deep into the soil. In contrast, a **fibrous root system** is located closer to the soil surface, and forms a dense network of roots that also helps prevent soil erosion (lawn grasses are a good example, as are wheat, rice, and corn). Some plants have a combination of tap roots and fibrous roots. Plants that grow in dry areas often have deep root systems, whereas plants growing in areas with abundant water are likely to have shallower root systems.

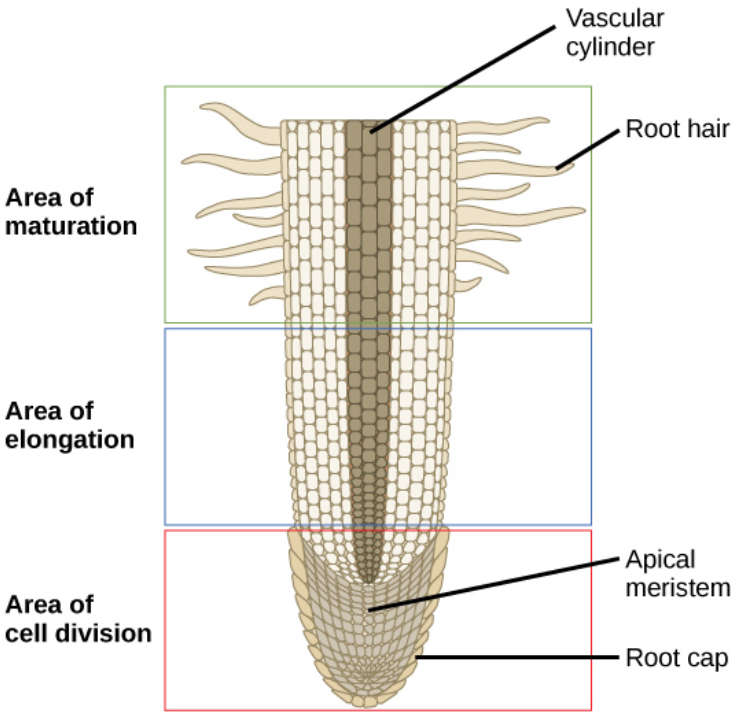
A longitudinal view of the root reveals the zones of cell division, elongation, and maturation. Cell division occurs in the apical meristem. Staining reveals different cell types in this light micrograph

of a wheat (*Triticum*) root cross section. Sclerenchyma cells of the exodermis and xylem cells stain red, and phloem cells stain blue. Other cell types stain black. The stele, or vascular tissue, is the area inside endodermis (indicated by a green ring). Root hairs are visible outside the epidermis. (credit: scale-bar data from Matt Russell) In (left) typical dicots, the vascular tissue forms an X shape in the center of the root. In (right) typical monocots, the phloem cells and the larger xylem cells form a characteristic ring around the central pith.

Root Growth and Anatomy

Root growth begins with seed germination. When the plant embryo emerges from the seed, the radicle of the embryo forms the root system. The tip of the root is protected by the **root cap**, a structure exclusive to roots and unlike any other plant structure. The root cap is continuously replaced because it gets damaged easily as the root pushes through soil. The root tip can be divided into three zones: a zone of cell division, a zone of elongation, and a zone of maturation and differentiation ([link]). The zone of cell division is closest to the root tip; it is made up of the actively dividing cells of the root meristem. The zone of elongation is where the newly formed cells increase in length, thereby lengthening the root. Beginning at the first root hair is the zone of cell maturation where the root cells begin to differentiate into special cell

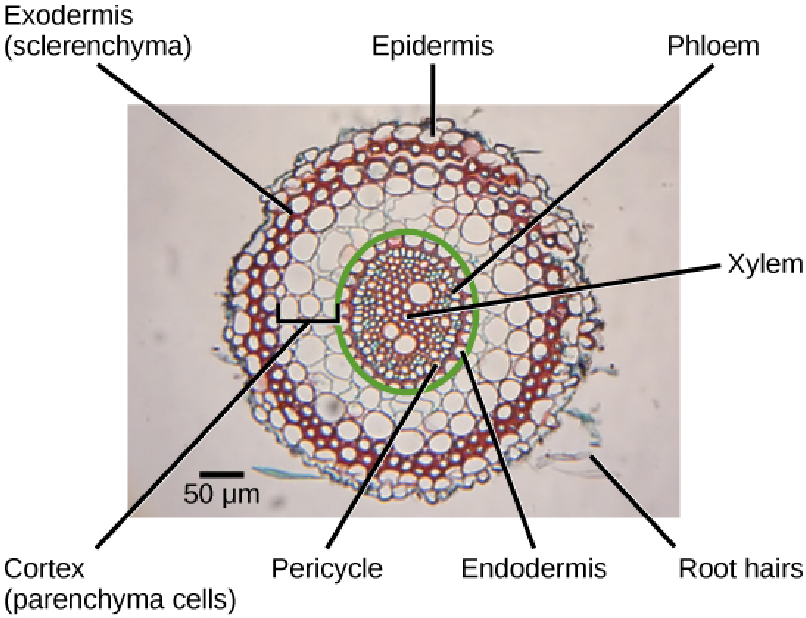
types. All three zones are in the first centimeter or so of the root tip.



The root has an outer layer of cells called the epidermis, which surrounds areas of ground tissue and vascular tissue. The epidermis provides protection and helps in absorption. **Root hairs**, which are extensions of root epidermal cells, increase the surface area of the root, greatly contributing to the absorption of water and minerals.

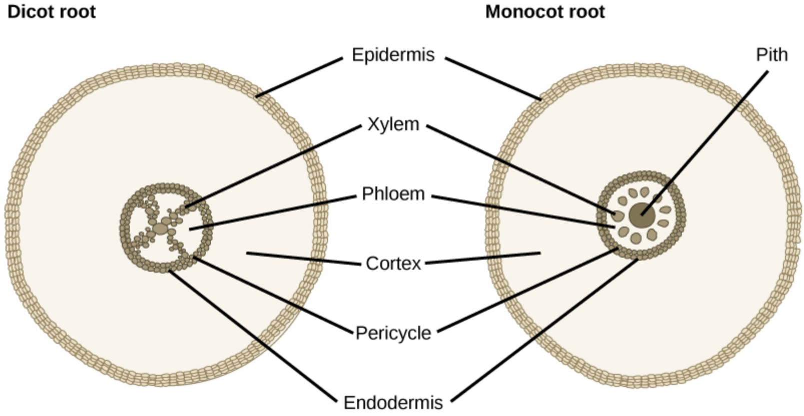
Inside the root, the ground tissue forms two regions: the cortex and the pith ([link]). Compared to stems, roots have lots of cortex and little pith. Both regions

include cells that store photosynthetic products. The cortex is between the epidermis and the vascular tissue, whereas the pith lies between the vascular tissue and the center of the root.



The vascular tissue in the root is arranged in the inner portion of the root, which is called the **stele** ([link]). A layer of cells known as the **endodermis** separates the stele from the ground tissue in the outer portion of the root. The endodermis is exclusive to roots, and serves as a checkpoint for materials entering the root's vascular system. A waxy substance called suberin is present on the walls of the endodermal cells. This waxy region, known as the **Casparian strip**, forces water and solutes to cross the plasma membranes of endodermal cells instead of slipping between the

cells. This ensures that only materials required by the root pass through the endodermis, while toxic substances and pathogens are generally excluded. The outermost cell layer of the root's vascular tissue is the **pericycle**, an area that can give rise to lateral roots. In dicot roots, the xylem and phloem of the stele are arranged alternately in an X shape, whereas in monocot roots, the vascular tissue is arranged in a ring around the pith.



Many vegetables are modified roots. The (a) banyan tree, also known as the strangler fig, begins life as an epiphyte in a host tree. Aerial roots extend to the ground and support the growing plant, which eventually strangles the host tree. The (b) screwpine develops aboveground roots that help support the plant in sandy soils. (credit a: modification of work by "psyberartist"/Flickr; credit b: modification of work by David Eikhoff)

Root Modifications

Root structures may be modified for specific

purposes. For example, some roots are bulbous and store starch. Aerial roots and prop roots are two forms of aboveground roots that provide additional support to anchor the plant. Tap roots, such as carrots, turnips, and beets, are examples of roots that are modified for food storage ([link]).

Epiphytic roots enable a plant to grow on another plant. For example, the epiphytic roots of orchids develop a spongy tissue to absorb moisture. The banyan tree (*Ficus* sp.) begins as an epiphyte, germinating in the branches of a host tree; aerial roots develop from the branches and eventually reach the ground, providing additional support ([link]). In screwpine (*Pandanus* sp.), a palm-like tree that grows in sandy tropical soils, aboveground prop roots develop from the nodes to provide additional support.

(a) (b)

Section Summary

Roots help to anchor a plant, absorb water and minerals, and serve as storage sites for food. Taproots and fibrous roots are the two main types of root systems. In a taproot system, a main root grows vertically downward with a few lateral roots. Fibrous root systems arise at the base of the stem, where a cluster of roots forms a dense network that is shallower than a taproot. The growing root tip is protected by a root cap. The root tip has three main zones: a zone of cell division (cells are actively dividing), a zone of elongation (cells increase in length), and a zone of maturation (cells differentiate to form different kinds of cells). Root vascular tissue conducts water, minerals, and sugars. In some habitats, the roots of certain plants may be modified to form aerial roots or epiphytic roots.

Review Questions

Roots that enable a plant to grow on another plant are called _____.

1. epiphytic roots
2. prop roots
3. adventitious roots
4. aerial roots

A

The _____ forces selective uptake of minerals in the root.

1. pericycle
2. epidermis
3. endodermis
4. root cap

C

Newly-formed root cells begin to form different cell types in the _____.

1. zone of elongation
2. zone of maturation
3. root meristem
4. zone of cell division

B

Critical Thinking Questions

Compare a tap root system with a fibrous root system. For each type, name a plant that provides a food in the human diet. Which type of root system is found in monocots? Which type of root system is found in dicots?

A tap root system has a single main root that grows down. A fibrous root system forms a dense network of roots that is closer to the soil surface. An example of a tap root system is a carrot. Grasses such as wheat, rice, and corn are examples of fibrous root systems. Fibrous root systems are found in monocots; tap root systems are found in dicots.

What might happen to a root if the pericycle disappeared?

The root would not be able to produce lateral roots.

Glossary

adventitious root
: aboveground root that arises from a plant part other than the radicle of the plant embryo

Casparian strip
: waxy coating that forces water to cross endodermal plasma membranes before entering the vascular cylinder, instead of moving between endodermal cells

endodermis
: layer of cells in the root that forms a selective barrier between the ground tissue and the vascular tissue, allowing water and minerals to enter the root while excluding toxins and pathogens

fibrous root system
: type of root system in which the roots arise from the base of the stem in a cluster, forming a dense network of roots; found in monocots

pericycle
: outer boundary of the stele from which lateral roots can arise

root cap
: protective cells covering the tip of the

growing root

root hair
: hair-like structure that is an extension of epidermal cells; increases the root surface area and aids in absorption of water and minerals

stele
: inner portion of the root containing the vascular tissue; surrounded by the endodermis

tap root system
: type of root system with a main root that grows vertically with few lateral roots; found in dicots

Leaves

By the end of this section, you will be able to do the following:

- Identify the parts of a typical leaf
- Describe the internal structure and function of a leaf
- Compare and contrast simple leaves and compound leaves
- List and describe examples of modified leaves

Leaves are the main sites for photosynthesis: the process by which plants synthesize food. Most leaves are usually green, due to the presence of chlorophyll in the leaf cells. However, some leaves may have different colors, caused by other plant pigments that mask the green chlorophyll.

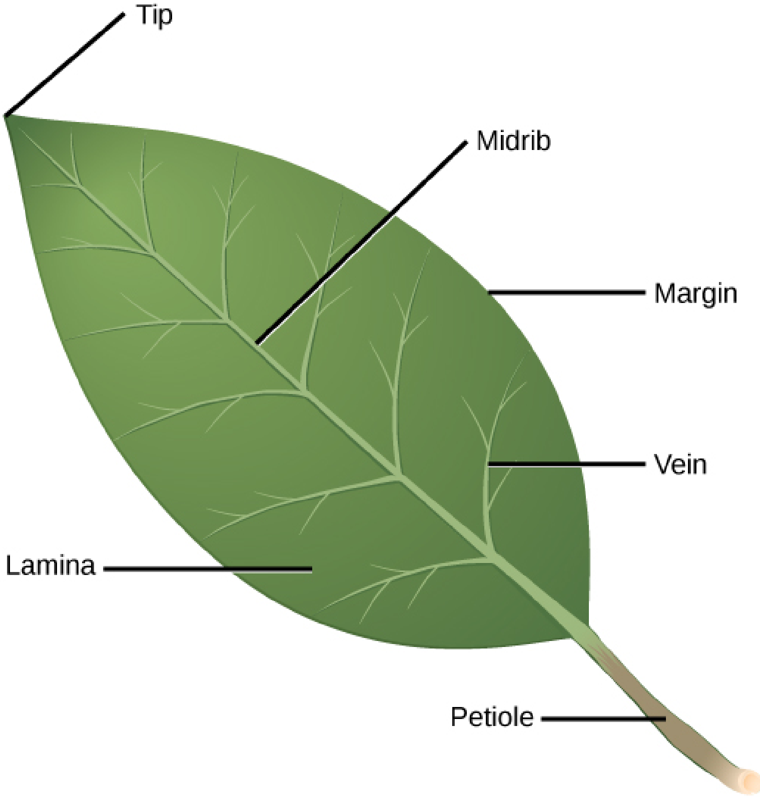
The thickness, shape, and size of leaves are adapted to the environment. Each variation helps a plant species maximize its chances of survival in a particular habitat. Usually, the leaves of plants growing in tropical rainforests have larger surface areas than those of plants growing in deserts or very cold conditions, which are likely to have a smaller surface area to minimize water loss.

Deceptively simple in appearance, a leaf is a highly efficient structure. (a) Tulip (*Tulipa*), a monocot, has leaves with parallel venation. The netlike venation in this (b) linden (*Tilia cordata*) leaf distinguishes it as a dicot. The (c) *Ginkgo biloba* tree has

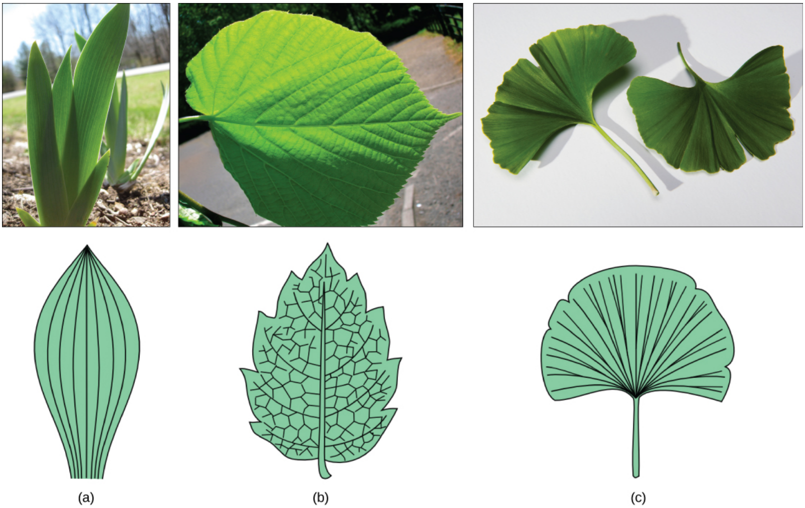
dichotomous venation. (credit a photo: modification of work by “Drewboy64”/Wikimedia Commons; credit b photo: modification of work by Roger Griffith; credit c photo: modification of work by "geishaboy500"/Flickr; credit abc illustrations: modification of work by Agnieszka Kwiecień)

Structure of a Typical Leaf

Each leaf typically has a leaf blade called the **lamina**, which is also the widest part of the leaf. Some leaves are attached to the plant stem by a **petiole**. Leaves that do not have a petiole and are directly attached to the plant stem are called **sessile** leaves. Small green appendages usually found at the base of the petiole are known as **stipules**. Most leaves have a midrib, which travels the length of the leaf and branches to each side to produce veins of vascular tissue. The edge of the leaf is called the margin. [link] shows the structure of a typical eudicot leaf.



Within each leaf, the vascular tissue forms veins. The arrangement of veins in a leaf is called the **venation** pattern. Monocots and dicots differ in their patterns of venation ([link]). Monocots have parallel venation; the veins run in straight lines across the length of the leaf without converging at a point. In dicots, however, the veins of the leaf have a net-like appearance, forming a pattern known as reticulate venation. One extant plant, the *Ginkgo biloba*, has dichotomous venation where the veins fork.



Leaf Arrangement

The arrangement of leaves on a stem is known as **phyllotaxy**. The number and placement of a plant’s leaves will vary depending on the species, with each species exhibiting a characteristic leaf arrangement. Leaves are classified as either alternate, spiral, or opposite. Plants that have only one leaf per node have leaves that are said to be either alternate—meaning the leaves alternate on each side of the stem in a flat plane—or spiral, meaning the leaves are arrayed in a spiral along the stem. In an opposite leaf arrangement, two leaves arise at the same point, with the leaves connecting opposite each other along the branch. If there are three or more leaves connected at a node, the leaf arrangement is classified as **whorled**.

Leaves may be simple or compound. In simple

leaves, the lamina is continuous. The (a) banana plant (*Musa* sp.) has simple leaves. In compound leaves, the lamina is separated into leaflets. Compound leaves may be palmate or pinnate. In (b) palmately compound leaves, such as those of the horse chestnut (*Aesculus hippocastanum*), the leaflets branch from the petiole. In (c) pinnately compound leaves, the leaflets branch from the midrib, as on a scrub hickory (*Carya floridana*). The (d) honey locust has double compound leaves, in which leaflets branch from the veins. (credit a: modification of work by "BazzaDaRambler"/Flickr; credit b: modification of work by Roberto Verzo; credit c: modification of work by Eric Dion; credit d: modification of work by Valerie Lykes)

Leaf Form

Leaves may be simple or compound ([link]). In a **simple leaf**, the blade is either completely undivided—as in the banana leaf—or it has lobes, but the separation does not reach the midrib, as in the maple leaf. In a **compound leaf**, the leaf blade is completely divided, forming leaflets, as in the locust tree. Each leaflet may have its own stalk, but is attached to the rachis. A **palmately compound leaf** resembles the palm of a hand, with leaflets radiating outwards from one point. Examples include the leaves of poison ivy, the buckeye tree, or the familiar houseplant *Schefflera* sp. (common name “umbrella plant”). **Pinnately compound**

leaves take their name from their feather-like appearance; the leaflets are arranged along the midrib, as in rose leaves (*Rosa* sp.), or the leaves of hickory, pecan, ash, or walnut trees.

(a) Simple

(b) Palmately compound

(c) Pinnately compound

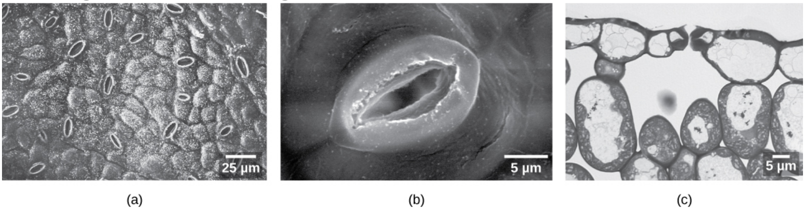
(d) Doubly compound

Visualized at 500x with a scanning electron microscope, several stomata are clearly visible on (a) the surface of this sumac (*Rhus glabra*) leaf. At 5,000x magnification, the guard cells of (b) a single stoma from lyre-leaved sand cress (*Arabidopsis lyrata)* have the appearance of lips that surround the opening. In this (c) light micrograph cross-section of an *A. lyrata* leaf, the guard cell pair is visible along with the large, sub-stomatal air space in the leaf. (credit: modification of work by Robert R. Wise;

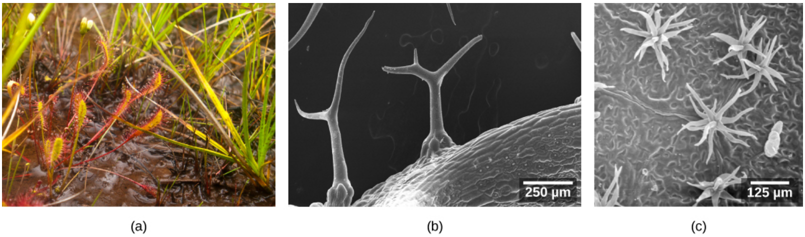
part c scale-bar data from Matt Russell) Trichomes give leaves a fuzzy appearance as in this (a) sundew (*Drosera* sp.). Leaf trichomes include (b) branched trichomes on the leaf of *Arabidopsis lyrata* and (c) multibranched trichomes on a mature *Quercus marilandica* leaf. (credit a: John Freeland; credit b, c: modification of work by Robert R. Wise; scale-bar data from Matt Russell)In the (a) leaf drawing, the central mesophyll is sandwiched between an upper and lower epidermis. The mesophyll has two layers: an upper palisade layer comprised of tightly packed, columnar cells, and a lower spongy layer, comprised of loosely packed, irregularly shaped cells. Stomata on the leaf underside allow gas exchange. A waxy cuticle covers all aerial surfaces of land plants to minimize water loss. These leaf layers are clearly visible in the (b) scanning electron micrograph. The numerous small bumps in the palisade parenchyma cells are chloroplasts. Chloroplasts are also present in the spongy parenchyma, but are not as obvious. The bumps protruding from the lower surface of the leave are glandular trichomes, which differ in structure from the stalked trichomes in [link]. (credit b: modification of work by Robert R. Wise) This scanning electron micrograph shows xylem and phloem in the leaf vascular bundle from the lyre-leaved sand cress (*Arabidopsis lyrata)*. (credit: modification of work by Robert R. Wise; scale-bar data from Matt Russell)

Leaf Structure and Function

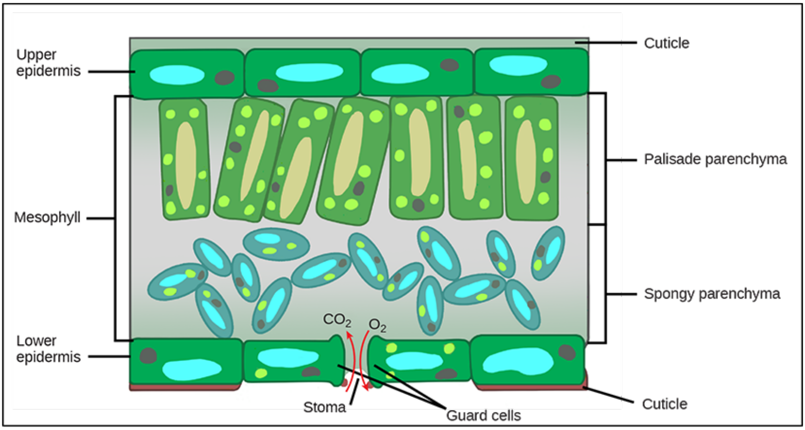
The outermost layer of the leaf is the epidermis; it is present on both sides of the leaf and is called the upper and lower epidermis, respectively. Botanists call the upper side the adaxial surface (or adaxis) and the lower side the abaxial surface (or abaxis). The epidermis helps in the regulation of gas exchange. It contains stomata ([link]): openings through which the exchange of gases takes place. Two guard cells surround each stoma, regulating its opening and closing.



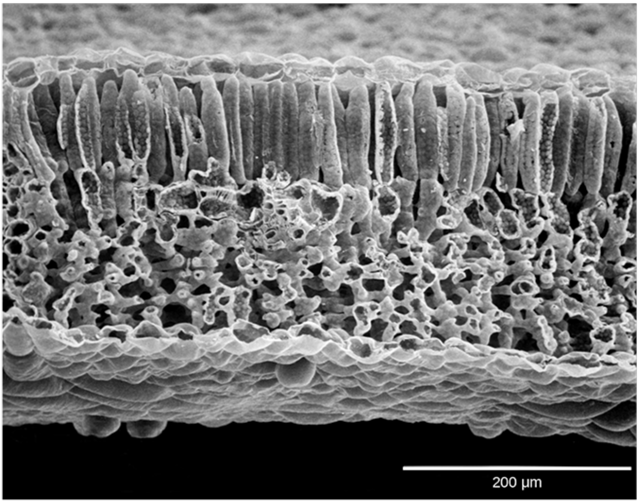
The epidermis is usually one cell layer thick; however, in plants that grow in very hot or very cold conditions, the epidermis may be several layers thick to protect against excessive water loss from transpiration. A waxy layer known as the **cuticle** covers the leaves of all plant species. The cuticle reduces the rate of water loss from the leaf surface. Other leaves may have small hairs (trichomes) on the leaf surface. Trichomes help to deter herbivory by restricting insect movements, or by storing toxic or bad-tasting compounds; they can also reduce the rate of transpiration by blocking air flow across the leaf surface ([link]).



Below the epidermis of dicot leaves are layers of cells known as the mesophyll, or “middle leaf.” The mesophyll of most leaves typically contains two arrangements of parenchyma cells: the palisade parenchyma and spongy parenchyma ([link]). The palisade parenchyma (also called the palisade mesophyll) has column-shaped, tightly packed cells, and may be present in one, two, or three layers. Below the palisade parenchyma are loosely arranged cells of an irregular shape. These are the cells of the spongy parenchyma (or spongy mesophyll). The air space found between the spongy parenchyma cells allows gaseous exchange between the leaf and the outside atmosphere through the stomata. In aquatic plants, the intercellular spaces in the spongy parenchyma help the leaf float. Both layers of the mesophyll contain many chloroplasts. Guard cells are the only epidermal cells to contain chloroplasts.

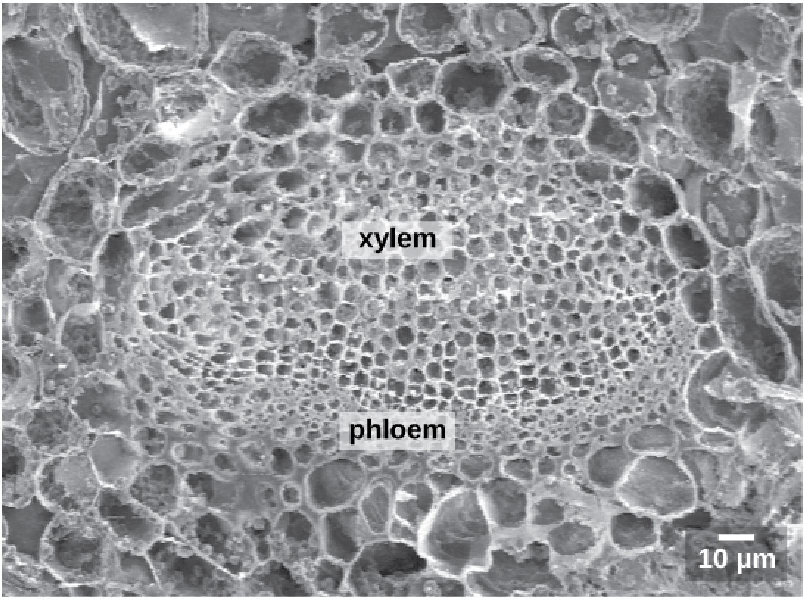


(a)



(b)

Like the stem, the leaf contains vascular bundles composed of xylem and phloem ([link]). The xylem consists of tracheids and vessels, which transport water and minerals to the leaves. The phloem transports the photosynthetic products from the leaf to the other parts of the plant. A single vascular bundle, no matter how large or small, always contains both xylem and phloem tissues.



Leaf Adaptations

Coniferous plant species that thrive in cold environments, like spruce, fir, and pine, have leaves that are reduced in size and needle-like in appearance. These needle-like leaves have sunken stomata and a smaller surface area: two attributes that aid in reducing water loss. In hot climates, plants such as cacti have leaves that are reduced to spines, which in combination with their succulent stems, help to conserve water. Many aquatic plants have leaves with wide lamina that can float on the surface of the water, and a thick waxy cuticle on the leaf surface that repels water.

Link to Learning

Watch “The Pale Pitcher Plant” episode of the video series *Plants Are Cool, Too,* a Botanical Society of America video about a carnivorous plant species found in Louisiana.

Evolution Connection

Plant Adaptations in Resource-Deficient Environments

Roots, stems, and leaves are structured to ensure that a plant can obtain the required sunlight, water, soil nutrients, and oxygen resources. Some remarkable adaptations have evolved to enable plant species to thrive in less than ideal habitats, where one or more of these resources is in short supply.

In tropical rainforests, light is often scarce, since many trees and plants grow close together and block much of the sunlight from reaching the forest floor. Many tropical plant species have exceptionally broad leaves to maximize the capture of sunlight. Other species are epiphytes: plants that grow on other plants that serve as a physical support. Such plants are able to grow high up in the canopy atop the branches of other trees, where sunlight is more plentiful. Epiphytes live on rain and minerals collected in the branches and leaves of the supporting plant. Bromeliads (members of the pineapple family), ferns, and orchids are

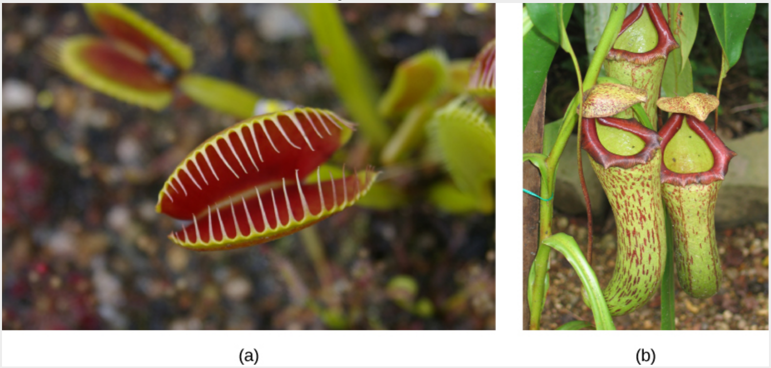
examples of tropical epiphytes ([link]). Many epiphytes have specialized tissues that enable them to efficiently capture and store water.

One of the most well known bromeliads is Spanish moss (*Tillandsia usneoides*), seen here in an oak tree. (credit: Kristine Paulus)

Some plants have special adaptations that help them to survive in nutrient-poor environments. Carnivorous plants, such as the Venus flytrap and the pitcher plant ([link]), grow in bogs where the soil is low in nitrogen. In these plants, leaves are

modified to capture insects. The insect-capturing leaves may have evolved to provide these plants with a supplementary source of much-needed nitrogen.

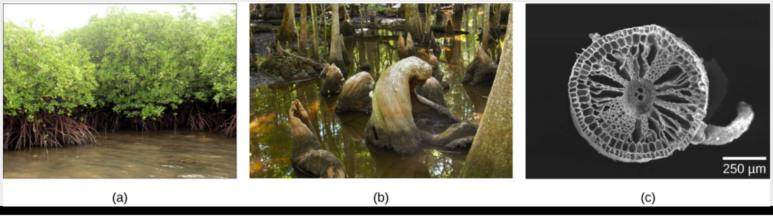
The (a) Venus flytrap has modified leaves that can capture insects. When an unlucky insect touches the trigger hairs inside the leaf, the trap suddenly closes. The opening of the (b) pitcher plant is lined with a slippery wax. Insects crawling on the lip slip and fall into a pool of water in the bottom of the pitcher, where they are digested by bacteria. The plant then absorbs the smaller molecules. (credit a: modification of work by Peter Shanks; credit b: modification of work by Tim Mansfield)



Many swamp plants have adaptations that enable them to thrive in wet areas, where their roots grow submerged underwater. In these aquatic areas, the soil is unstable and little oxygen is available to reach the roots. Trees such as mangroves (*Rhizophora* sp.) growing in coastal waters produce aboveground roots that help support the tree ([link]). Some species of mangroves, as well as cypress trees, have pneumatophores: upward-

growing roots containing pores and pockets of tissue specialized for gas exchange. Wild rice is an aquatic plant with large air spaces in the root cortex. The air-filled tissue—called aerenchyma—provides a path for oxygen to diffuse down to the root tips, which are embedded in oxygen-poor bottom sediments.

The branches of (a) mangrove trees develop aerial roots, which descend to the ground and help to anchor the trees. (b) Cypress trees and some mangrove species have upward-growing roots called pneumatophores that are involved in gas exchange. Aquatic plants such as (c) wild rice have large spaces in the root cortex called aerenchyma, visualized here using scanning electron microscopy. (credit a: modification of work by Roberto Verzo; credit b: modification of work by Duane Burdick; credit c: modification of work by Robert R. Wise)



Link to Learning

Watch *Venus Flytraps: Jaws of Death*, an extraordinary BBC close-up of the Venus flytrap in action.

https://www.openstax.org/l/venus_flytrap

Section Summary

Leaves are the main site of photosynthesis. A typical leaf consists of a lamina (the broad part of the leaf, also called the blade) and a petiole (the stalk that attaches the leaf to a stem). The arrangement of leaves on a stem, known as phyllotaxy, enables maximum exposure to sunlight. Each plant species has a characteristic leaf arrangement and form. The pattern of leaf arrangement may be alternate, opposite, or spiral, while leaf form may be simple or compound. Leaf tissue consists of the epidermis, which forms the outermost cell layer, and mesophyll and vascular tissue, which make up the inner portion of the leaf. In some plant species, leaf form is modified to form structures such as tendrils, spines, bud scales, and needles.

Review Questions

The stalk of a leaf is known as the _____.

1. petiole
2. lamina
3. stipule
4. rachis

A

Leaflets are a characteristic of ______ leaves.

1. alternate
2. whorled
3. compound
4. opposite

C

Cells of the ______ contain chloroplasts.

1. epidermis
2. vascular tissue
3. stomata
4. mesophyll

D

Which of the following is most likely to be

found in a desert environment?

1. broad leaves to capture sunlight
2. spines instead of leaves
3. needle-like leaves
4. wide, flat leaves that can float

B

Critical Thinking Questions

How do dicots differ from monocots in terms of leaf structure?

Monocots have leaves with parallel venation, and dicots have leaves with reticulate, net-like venation.

Describe an example of a plant with leaves that are adapted to cold temperatures.

Conifers such as spruce, fir, and pine have needle-shaped leaves with sunken stomata, helping to reduce water loss.

Glossary

compound leaf
: leaf in which the leaf blade is subdivided to form leaflets, all attached to the midrib

cuticle
: waxy protective layer on the leaf surface

lamina
: leaf blade

palmately compound leaf
: leaf type with leaflets that emerge from a point, resembling the palm of a hand

petiole
: stalk of the leaf

phyllotaxy
: arrangement of leaves on a stem

pinnately compound leaf
: leaf type with a divided leaf blade consisting of leaflets arranged on both sides of the midrib

sessile
: leaf without a petiole that is attached directly to the plant stem

simple leaf

leaf type in which the lamina is completely undivided or merely lobed

stipule
: small green structure found on either side of the leaf stalk or petiole

venation
: pattern of veins in a leaf; may be parallel (as in monocots), reticulate (as in dicots), or dichotomous (as in *Gingko biloba*)

whorled
: pattern of leaf arrangement in which three or more leaves are connected at a node

Transport of Water and Solutes in Plants

By the end of this section, you will be able to do the following:

- Define water potential and explain how it is influenced by solutes, pressure, gravity, and the matric potential
- Describe how water potential, evapotranspiration, and stomatal regulation influence how water is transported in plants
- Explain how photosynthates are transported in plants

The structure of plant roots, stems, and leaves facilitates the transport of water, nutrients, and photosynthates throughout the plant. The phloem and xylem are the main tissues responsible for this movement. Water potential, evapotranspiration, and stomatal regulation influence how water and nutrients are transported in plants. To understand how these processes work, we must first understand the energetics of water potential.

With heights nearing 116 meters, (a) coastal redwoods (*Sequoia sempervirens*) are the tallest trees in the world. Plant roots can easily generate enough force to (b) buckle and break concrete sidewalks, much to the dismay of homeowners and city maintenance departments. (credit a: modification of work by Bernt Rostad; credit b: modification of work by Pedestrians Educating Drivers on Safety, Inc.)When (a) total water potential (Ψ_{total}) is lower

outside the cells than inside, water moves out of the cells and the plant wilts. When (b) the total water potential is higher outside the plant cells than inside, water moves into the cells, resulting in turgor pressure (Ψ_p) and keeping the plant erect. (credit: modification of work by Victor M. Vicente Selvas)

Water Potential

Plants are phenomenal hydraulic engineers. Using only the basic laws of physics and the simple manipulation of potential energy, plants can move water to the top of a 116-meter-tall tree ([link]**a**). Plants can also use hydraulics to generate enough force to split rocks and buckle sidewalks ([link]**b**). Plants achieve this because of water potential.

(a)

(b)

Water potential is a measure of the potential energy in water. Plant physiologists are not interested in the energy in any one particular

aqueous system, but are very interested in water movement between two systems. In practical terms, therefore, water potential is the difference in potential energy between a given water sample and pure water (at atmospheric pressure and ambient temperature). Water potential is denoted by the Greek letter ψ (*psi*) and is expressed in units of pressure (pressure is a form of energy) called **megapascals** (MPa). The potential of pure water ($\Psi_{wpure\ H2O}$) is, by convenience of definition, designated a value of zero (even though pure water contains plenty of potential energy, that energy is ignored). Water potential values for the water in a plant root, stem, or leaf are therefore expressed relative to $\Psi_{wpure\ H2O}$.

The water potential in plant solutions is influenced by solute concentration, pressure, gravity, and factors called matrix effects. Water potential can be broken down into its individual components using the following equation:

$$\Psi \text{ system} = \Psi \text{ total} = \Psi \text{ s} + \Psi \text{ p} + \Psi \text{ g} + \Psi \text{ m}$$

where Ψ_s, Ψ_p, Ψ_g, and Ψ_m refer to the solute, pressure, gravity, and matric potentials, respectively. “System” can refer to the water potential of the soil water (Ψ_{soil}), root water (Ψ_{root}), stem water (Ψ_{stem}), leaf water (Ψ_{leaf}) or the water in the atmosphere ($\Psi_{atmosphere}$): whichever aqueous system is under consideration. As the individual components change, they raise or lower the total

water potential of a system. When this happens, water moves to equilibrate, moving from the system or compartment with a higher water potential to the system or compartment with a lower water potential. This brings the difference in water potential between the two systems ($\Delta\Psi$) back to zero ($\Delta\Psi = 0$). Therefore, for water to move through the plant from the soil to the air (a process called transpiration), Ψ_{soil} must be $> \Psi_{root} > \Psi_{stem} > \Psi_{leaf} > \Psi_{atmosphere}$.

Water only moves in response to $\Delta\Psi$, not in response to the individual components. However, because the individual components influence the total Ψ_{system}, by manipulating the individual components (especially Ψ_s), a plant can control water movement.

Solute Potential

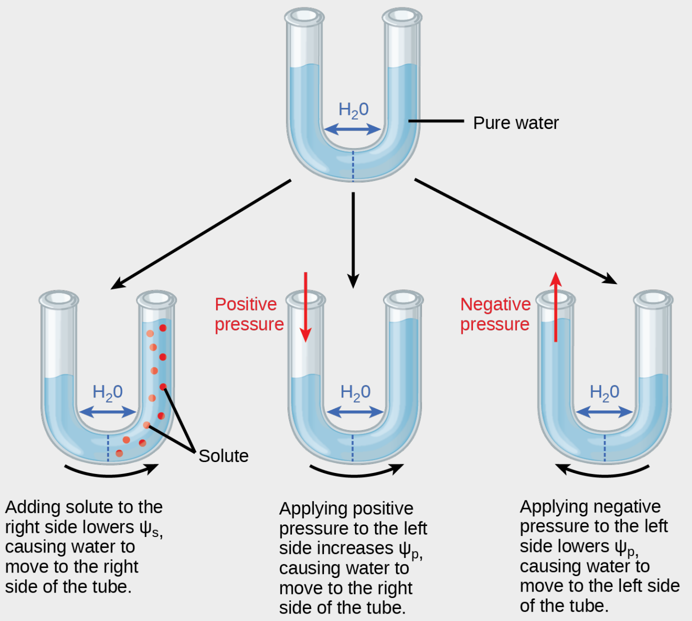
Solute potential (Ψ_s), also called osmotic potential, is related to the solute concentration (in molarity). That relationship is given by the van 't Hoff equation: $\Psi_s = -Mi\,RT$; where M is the molar concentration of the solute, *i* is the van 't Hoff factor (the ratio of the amount of particles in the solution to amount of formula units dissolved), R is the ideal gas constant, and T is temperature in Kelvin degrees. The solute potential is negative in a plant cell and zero in distilled water. Typical values for cell cytoplasm are –0.5 to –1.0 MPa. Solutes reduce

water potential (resulting in a negative Ψ_w) by consuming some of the potential energy available in the water. Solute molecules can dissolve in water because water molecules can bind to them via hydrogen bonds; a hydrophobic molecule like oil, which cannot bind to water, cannot go into solution. The energy in the hydrogen bonds between solute molecules and water is no longer available to do work in the system because it is tied up in the bond. In other words, the amount of available potential energy is reduced when solutes are added to an aqueous system. Thus, Ψ_s decreases with increasing solute concentration. Because Ψ_s is one of the four components of Ψ_{system} or Ψ_{total}, a decrease in Ψ_s will cause a decrease in Ψ_{total}. The internal water potential of a plant cell is more negative than pure water because of the cytoplasm's high solute content ([link]). Because of this difference in water potential water will move from the soil into a plant's root cells via the process of osmosis. This is why solute potential is sometimes called osmotic potential.

Plant cells can metabolically manipulate Ψ_s (and by extension, Ψ_{total}) by adding or removing solute molecules. Therefore, plants have control over Ψ_{total} via their ability to exert metabolic control over Ψ_s.

In this example with a semipermeable membrane between two aqueous systems, water will move

from a region of higher to lower water potential until equilibrium is reached. Solutes (Ψ_s), pressure (Ψ_p), and gravity (Ψ_g) influence total water potential for each side of the tube ($\Psi_{total\ right\ or\ left}$), and therefore, the difference between Ψ_{total} on each side ($\Delta\Psi$). (Ψ_m , the potential due to interaction of water with solid substrates, is ignored in this example because glass is not especially hydrophilic). Water moves in response to the difference in water potential between two systems (the left and right sides of the tube).



Positive water potential is placed on the left side of the tube by increasing Ψ_p such that the water level rises on the right side. Could you equalize the water level on each side of the tube by adding solute, and if so, how?

Pressure Potential

Pressure potential (Ψ_p), also called turgor potential, may be positive or negative ([link]). Because pressure is an expression of energy, the higher the pressure, the more potential energy in a system, and vice versa. Therefore, a positive Ψp (compression) increases Ψ_{total}, and a negative Ψ_p (tension) decreases Ψ_{total}. Positive pressure inside cells is contained by the cell wall, producing turgor pressure. Pressure potentials are typically around 0.6–0.8 MPa, but can reach as high as 1.5 MPa in a well-watered plant. A Ψ_p of 1.5 MPa equates to 210 pounds per square inch (1.5 MPa x 140 lb/in^{-2} MPa^{-1} = 210 lb/in^{-2}). As a comparison, most automobile tires are kept at a pressure of 30–34 psi. An example of the effect of turgor pressure is the wilting of leaves and their restoration after the plant has been watered ([link]). Water is lost from the leaves via transpiration (approaching Ψ_p = 0 MPa at the wilting point) and restored by uptake via the roots.

A plant can manipulate Ψ_p via its ability to manipulate Ψ_s and by the process of osmosis. If a plant cell increases the cytoplasmic solute concentration, Ψ_s will decline, Ψ_{total} will decline, the $\Delta\Psi$ between the cell and the surrounding tissue will decline, water will move into the cell by osmosis, and Ψ_p will increase. Ψ_p is also under indirect plant control via the opening and closing of

stomata. Stomatal openings allow water to evaporate from the leaf, reducing Ψ_p and Ψ_{total} of the leaf and increasing Ψ between the water in the leaf and the petiole, thereby allowing water to flow from the petiole into the leaf.

(a) (b)

Gravity Potential

Gravity potential (Ψ_g) is always negative to zero in a plant with no height. It always removes or consumes potential energy from the system. The force of gravity pulls water downwards to the soil, reducing the total amount of potential energy in the water in the plant (Ψ_{total}). The taller the plant, the taller the water column, and the more influential Ψ_g becomes. On a cellular scale and in short plants, this effect is negligible and easily ignored. However, over the height of a tall tree like a giant coastal redwood, the gravitational pull of –0.1 MPa m^{-1} is equivalent to an extra 1 MPa of resistance that must be overcome for water to reach the leaves of the tallest trees. Plants are unable to manipulate Ψ_g.

Matric Potential

Matric potential (Ψ_m) is always negative to zero. In a dry system, it can be as low as –2 MPa in a dry seed, and it is zero in a water-saturated system. The binding of water to a matrix always removes or consumes potential energy from the system. Ψ_m is similar to solute potential because it involves tying up the energy in an aqueous system by forming hydrogen bonds between the water and some other component. However, in solute potential, the other components are soluble, hydrophilic solute molecules, whereas in Ψ_m, the other components are insoluble, hydrophilic molecules of the plant cell wall. Every plant cell has a cellulosic cell wall and the cellulose in the cell walls is hydrophilic, producing a matrix for adhesion of water: hence the name matric potential. Ψ_m is very large (negative) in dry tissues such as seeds or drought-affected soils. However, it quickly goes to zero as the seed takes up water or the soil hydrates. Ψ_m cannot be manipulated by the plant and is typically ignored in well-watered roots, stems, and leaves.

Plants are suited to their local environment. (a) Xerophytes, like this prickly pear cactus (*Opuntia sp.*) and (b) epiphytes such as this tropical *Aeschynanthus perrottetii* have adapted to very limited water resources. The leaves of a prickly pear are modified into spines, which lowers the surface-to-volume ratio and reduces water loss. Photosynthesis takes place in the stem, which also

stores water. (b) *A. perottetii* leaves have a waxy cuticle that prevents water loss. (c) Goldenrod (*Solidago sp.*) is a mesophyte, well suited for moderate environments. (d) Hydrophytes, like this fragrant water lily (*Nymphaea odorata*), are adapted to thrive in aquatic environments. (credit a: modification of work by Jon Sullivan; credit b: modification of work by L. Shyamal/Wikimedia Commons; credit c: modification of work by Huw Williams; credit d: modification of work by Jason Hollinger)

Movement of Water and Minerals in the Xylem

Solutes, pressure, gravity, and matric potential are all important for the transport of water in plants. Water moves from an area of higher total water potential (higher Gibbs free energy) to an area of lower total water potential. Gibbs free energy is the energy associated with a chemical reaction that can be used to do work. This is expressed as $\Delta\Psi$.

Transpiration is the loss of water from the plant through evaporation at the leaf surface. It is the main driver of water movement in the xylem. Transpiration is caused by the evaporation of water at the leaf–atmosphere interface; it creates negative pressure (tension) equivalent to –2 MPa at the leaf surface. This value varies greatly depending on the vapor pressure deficit, which can be negligible at

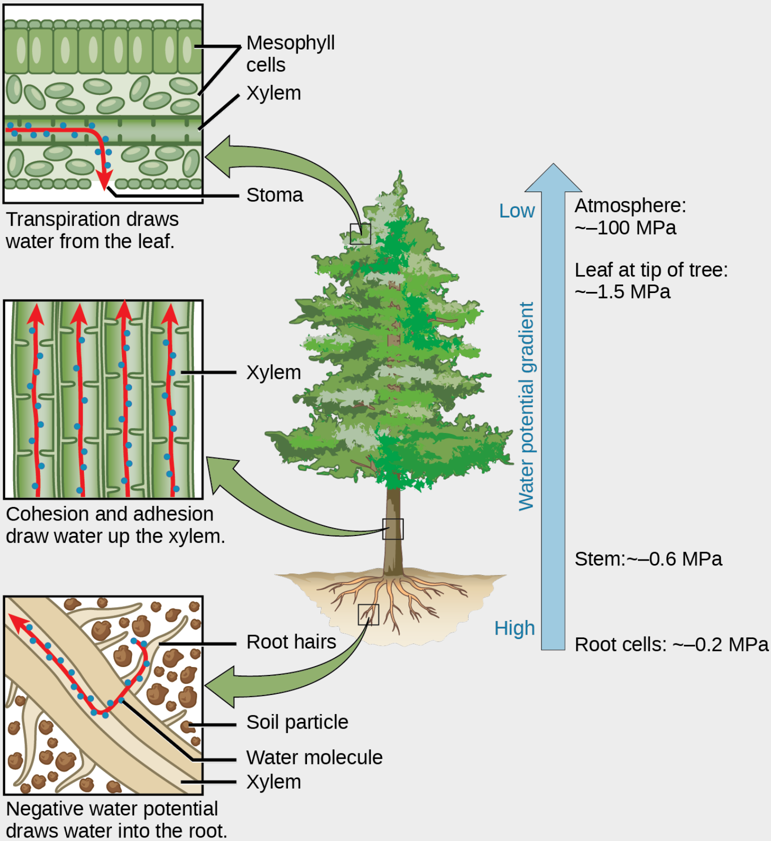
high relative humidity (RH) and substantial at low RH. Water from the roots is pulled up by this tension. At night, when stomata shut and transpiration stops, the water is held in the stem and leaf by the adhesion of water to the cell walls of the xylem vessels and tracheids, and the cohesion of water molecules to each other. This is called the cohesion–tension theory of sap ascent.

Inside the leaf at the cellular level, water on the surface of mesophyll cells saturates the cellulose microfibrils of the primary cell wall. The leaf contains many large intercellular air spaces for the exchange of oxygen for carbon dioxide, which is required for photosynthesis. The wet cell wall is exposed to this leaf internal air space, and the water on the surface of the cells evaporates into the air spaces, decreasing the thin film on the surface of the mesophyll cells. This decrease creates a greater tension on the water in the mesophyll cells ([link]), thereby increasing the pull on the water in the xylem vessels. The xylem vessels and tracheids are structurally adapted to cope with large changes in pressure. Rings in the vessels maintain their tubular shape, much like the rings on a vacuum cleaner hose keep the hose open while it is under pressure. Small perforations between vessel elements reduce the number and size of gas bubbles that can form via a process called cavitation. The formation of gas bubbles in xylem interrupts the continuous stream of water from the base to the top of the plant,

causing a break termed an embolism in the flow of xylem sap. The taller the tree, the greater the tension forces needed to pull water, and the more cavitation events. In larger trees, the resulting embolisms can plug xylem vessels, making them nonfunctional.

Visual Connection

The cohesion–tension theory of sap ascent is shown. Evaporation from the mesophyll cells produces a negative water potential gradient that causes water to move upwards from the roots through the xylem.



Which of the following statements is false?

1. Negative water potential draws water into the root hairs. Cohesion and adhesion draw water up the xylem. Transpiration draws water from the leaf.
2. Negative water potential draws water into the root hairs. Cohesion and adhesion draw water up the phloem. Transpiration draws water from the leaf.
3. Water potential decreases from the roots to the top of the plant.
4. Water enters the plants through root hairs and

exits through stoma.

Transpiration—the loss of water vapor to the atmosphere through stomata—is a passive process, meaning that metabolic energy in the form of ATP is not required for water movement. The energy driving transpiration is the difference in energy between the water in the soil and the water in the atmosphere. However, transpiration is tightly controlled.

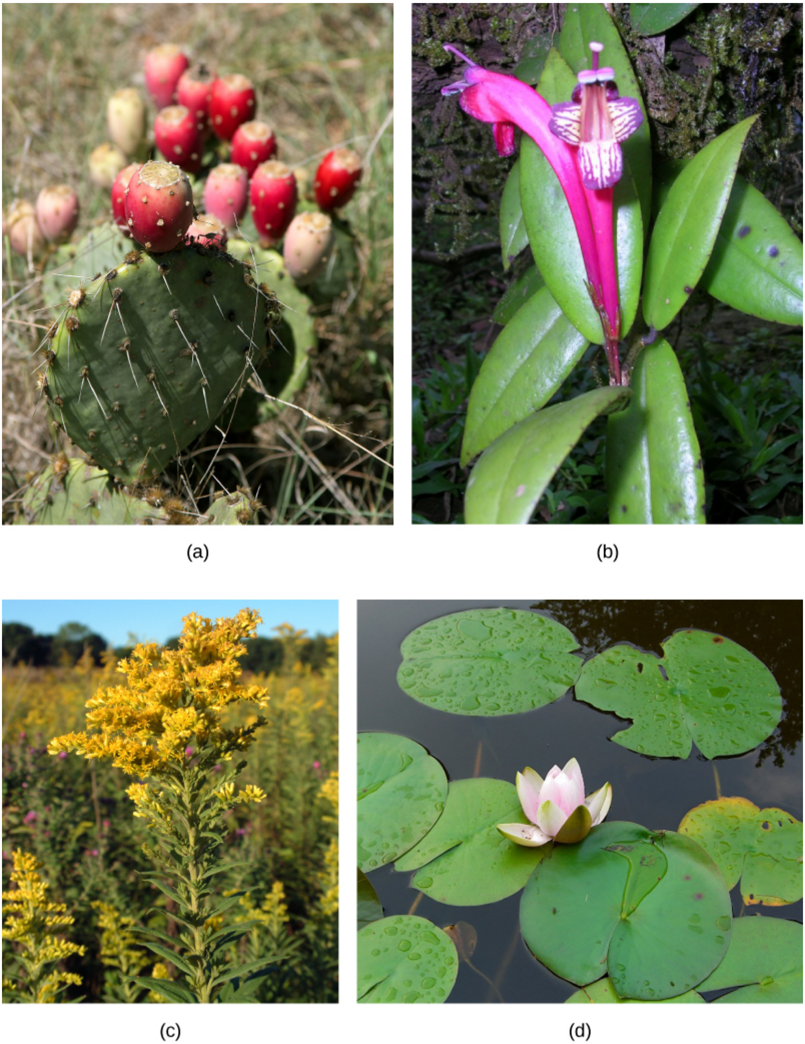
Control of Transpiration

The atmosphere to which the leaf is exposed drives transpiration, but also causes massive water loss from the plant. Up to 90 percent of the water taken up by roots may be lost through transpiration.

Leaves are covered by a waxy **cuticle** on the outer surface that prevents the loss of water. Regulation of transpiration, therefore, is achieved primarily through the opening and closing of stomata on the leaf surface. Stomata are surrounded by two specialized cells called guard cells, which open and close in response to environmental cues such as light intensity and quality, leaf water status, and carbon dioxide concentrations. Stomata must open to allow air containing carbon dioxide and oxygen to diffuse into the leaf for photosynthesis and

respiration. When stomata are open, however, water vapor is lost to the external environment, increasing the rate of transpiration. Therefore, plants must maintain a balance between efficient photosynthesis and water loss.

Plants have evolved over time to adapt to their local environment and reduce transpiration ([link]). Desert plant (xerophytes) and plants that grow on other plants (epiphytes) have limited access to water. Such plants usually have a much thicker waxy cuticle than those growing in more moderate, well-watered environments (mesophytes). Aquatic plants (hydrophytes) also have their own set of anatomical and morphological leaf adaptations.



Xerophytes and epiphytes often have a thick covering of trichomes or of stomata that are sunken below the leaf's surface. Trichomes are specialized hair-like epidermal cells that secrete oils and substances. These adaptations impede air flow across the stomatal pore and reduce transpiration. Multiple epidermal layers are also commonly found

in these types of plants.

Phloem is comprised of cells called sieve-tube elements. Phloem sap travels through perforations called sieve tube plates. Neighboring companion cells carry out metabolic functions for the sieve-tube elements and provide them with energy. Lateral sieve areas connect the sieve-tube elements to the companion cells.Sucrose is actively transported from source cells into companion cells and then into the sieve-tube elements. This reduces the water potential, which causes water to enter the phloem from the xylem. The resulting positive pressure forces the sucrose-water mixture down toward the roots, where sucrose is unloaded. Transpiration causes water to return to the leaves through the xylem vessels.

Transportation of Photosynthates in the Phloem

Plants need an energy source to grow. In seeds and bulbs, food is stored in polymers (such as starch) that are converted by metabolic processes into sucrose for newly developing plants. Once green shoots and leaves are growing, plants are able to produce their own food by photosynthesizing. The products of photosynthesis are called photosynthates, which are usually in the form of simple sugars such as sucrose.

Structures that produce photosynthates for the

growing plant are referred to as **sources**. Sugars produced in sources, such as leaves, need to be delivered to growing parts of the plant via the phloem in a process called **translocation**. The points of sugar delivery, such as roots, young shoots, and developing seeds, are called **sinks**. Seeds, tubers, and bulbs can be either a source or a sink, depending on the plant's stage of development and the season.

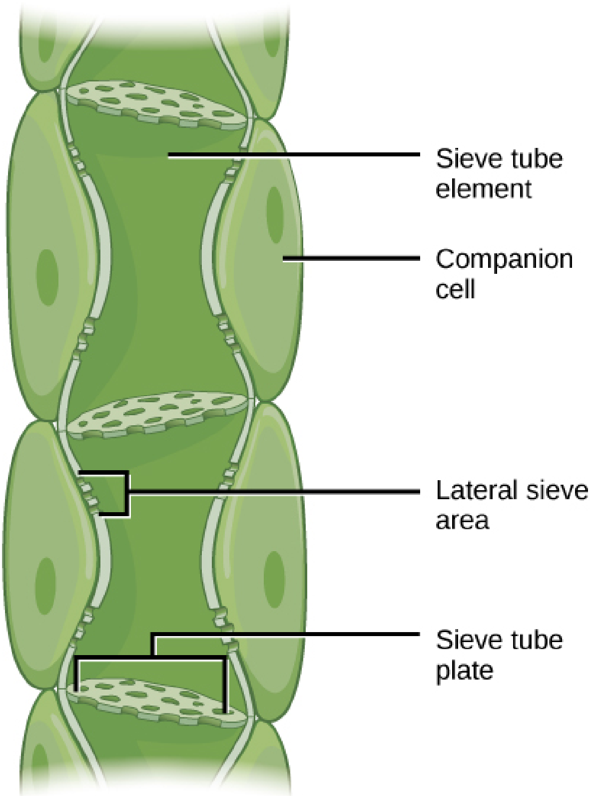
The products from the source are usually translocated to the nearest sink through the phloem. For example, the highest leaves will send photosynthates upward to the growing shoot tip, whereas lower leaves will direct photosynthates downward to the roots. Intermediate leaves will send products in both directions, unlike the flow in the xylem, which is always unidirectional (soil to leaf to atmosphere). The pattern of photosynthate flow changes as the plant grows and develops. Photosynthates are directed primarily to the roots early on, to shoots and leaves during vegetative growth, and to seeds and fruits during reproductive development. They are also directed to tubers for storage.

Translocation: Transport from Source to Sink

Photosynthates, such as sucrose, are produced in the mesophyll cells of photosynthesizing leaves. From there they are translocated through the phloem to

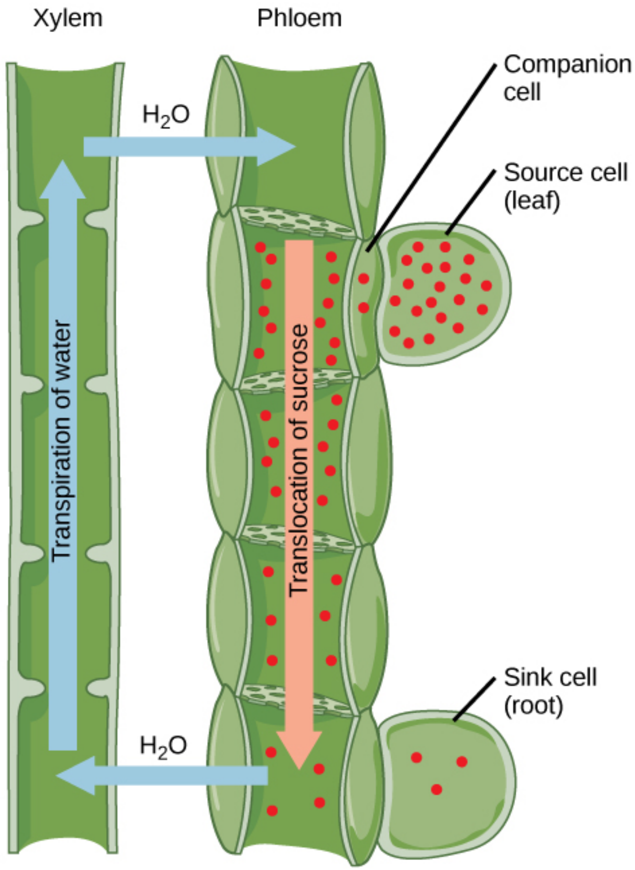
where they are used or stored. Mesophyll cells are connected by cytoplasmic channels called plasmodesmata. Photosynthates move through these channels to reach phloem sieve-tube elements (STEs) in the vascular bundles. From the mesophyll cells, the photosynthates are loaded into the phloem STEs. The sucrose is actively transported against its concentration gradient (a process requiring ATP) into the phloem cells using the electrochemical potential of the proton gradient. This is coupled to the uptake of sucrose with a carrier protein called the sucrose-H^+ symporter.

Phloem STEs have reduced cytoplasmic contents, and are connected by a sieve plate with pores that allow for pressure-driven bulk flow, or translocation, of phloem sap. Companion cells are associated with STEs. They assist with metabolic activities and produce energy for the STEs ([link]).



Once in the phloem, the photosynthates are translocated to the closest sink. Phloem sap is an aqueous solution that contains up to 30 percent sugar, minerals, amino acids, and plant growth regulators. The high percentage of sugar decreases Ψ_s, which decreases the total water potential and causes water to move by osmosis from the adjacent xylem into the phloem tubes, thereby increasing pressure. This increase in total water potential causes the bulk flow of phloem from source to sink ([link]). Sucrose concentration in the sink cells is lower than in the phloem STEs because the sink

sucrose has been metabolized for growth, or converted to starch for storage or other polymers, such as cellulose, for structural integrity. Unloading at the sink end of the phloem tube occurs by either diffusion or active transport of sucrose molecules from an area of high concentration to one of low concentration. Water diffuses from the phloem by osmosis and is then transpired or recycled via the xylem back into the phloem sap.



Section Summary

Water potential (Ψ) is a measure of the difference in potential energy between a water sample and pure water. The water potential in plant solutions is influenced by solute concentration, pressure, gravity, and matric potential. Water potential and transpiration influence how water is transported through the xylem in plants. These processes are regulated by stomatal opening and closing. Photosynthates (mainly sucrose) move from sources to sinks through the plant's phloem. Sucrose is actively loaded into the sieve-tube elements of the phloem. The increased solute concentration causes water to move by osmosis from the xylem into the phloem. The positive pressure that is produced pushes water and solutes down the pressure gradient. The sucrose is unloaded into the sink, and the water returns to the xylem vessels.

Visual Connection Questions

[link] Positive water potential is placed on the left side of the tube by increasing Ψ_p such that the water level rises on the right side. Could you equalize the water level on each side of the tube by adding solute, and if so, how?

[link] Yes, you can equalize the water level by adding the solute to the left side of the tube such that water moves toward the left until the water levels are equal.

[link] Which of the following statements is false?

1. Negative water potential draws water into the root hairs. Cohesion and adhesion draw water up the xylem. Transpiration draws water from the leaf.
2. Negative water potential draws water into the root hairs. Cohesion and adhesion draw water up the phloem. Transpiration draws water from the leaf.
3. Water potential decreases from the roots to the top of the plant.
4. Water enters the plants through root hairs and exits through stoma.

[link] B.

Review Questions

When stomata open, what occurs?

1. Water vapor is lost to the external environment, increasing the rate of transpiration.
2. Water vapor is lost to the external environment, decreasing the rate of transpiration.
3. Water vapor enters the spaces in the mesophyll, increasing the rate of transpiration.
4. Water vapor enters the spaces in the mesophyll, decreasing the rate of transpiration.

A

Which cells are responsible for the movement of photosynthates through a plant?

1. tracheids, vessel elements
2. tracheids, companion cells
3. vessel elements, companion cells
4. sieve-tube elements, companion cells

D

Critical Thinking Questions

The process of bulk flow transports fluids in a plant. Describe the two main bulk flow processes.

The process of bulk flow moves water up the xylem and moves photosynthates (solutes) up and down the phloem.

Glossary

cuticle
: waxy covering on the outside of the leaf and stem that prevents the loss of water

megapascal (MPa)
: pressure units that measure water potential

sink
: growing parts of a plant, such as roots and young leaves, which require photosynthate

source
: organ that produces photosynthate for a plant

translocation
: mass transport of photosynthates from source to sink in vascular plants

transpiration
: loss of water vapor to the atmosphere through

stomata

water potential (Ψ_w)
: the potential energy of a water solution per unit volume in relation to pure water at atmospheric pressure and ambient temperature

Plant Sensory Systems and Responses

By the end of this section, you will be able to do the following:

- Describe how red and blue light affect plant growth and metabolic activities
- Discuss gravitropism
- Understand how hormones affect plant growth and development
- Describe thigmotropism, thigmonastism, and thigmogenesis
- Explain how plants defend themselves from predators and respond to wounds

Animals can respond to environmental factors by moving to a new location. Plants, however, are rooted in place and must respond to the surrounding environmental factors. Plants have sophisticated systems to detect and respond to light, gravity, temperature, and physical touch. Receptors sense environmental factors and relay the information to effector systems—often through intermediate chemical messengers—to bring about plant responses.

The biologically inactive form of phytochrome (Pr) is converted to the biologically active form Pfr under illumination with red light. Far-red light and darkness convert the molecule back to the inactive form. Azure bluets (*Houstonia caerulea*) display a phototropic response by bending toward the light. (credit: Cory Zanker)

Plant Responses to Light

Plants have a number of sophisticated uses for light that go far beyond their ability to photosynthesize low-molecular-weight sugars using only carbon dioxide, light, and water. **Photomorphogenesis** is the growth and development of plants in response to light. It allows plants to optimize their use of light and space. **Photoperiodism** is the ability to use light to track time. Plants can tell the time of day and time of year by sensing and using various wavelengths of sunlight. **Phototropism** is a directional response that allows plants to grow towards, or even away from, light.

The sensing of light in the environment is important to plants; it can be crucial for competition and survival. The response of plants to light is mediated by different photoreceptors, which are comprised of a protein covalently bonded to a light-absorbing pigment called a **chromophore**. Together, the two are called a chromoprotein.

The red/far-red and violet-blue regions of the visible light spectrum trigger structural development in plants. Sensory photoreceptors absorb light in these particular regions of the visible light spectrum because of the quality of light available in the daylight spectrum. In terrestrial habitats, light absorption by chlorophylls peaks in the blue and red regions of the spectrum. As light filters through the

canopy and the blue and red wavelengths are absorbed, the spectrum shifts to the far-red end, shifting the plant community to those plants better adapted to respond to far-red light. Blue-light receptors allow plants to gauge the direction and abundance of sunlight, which is rich in blue–green emissions. Water absorbs red light, which makes the detection of blue light essential for algae and aquatic plants.

The Phytochrome System and the Red/Far-Red Response

The **phytochromes** are a family of chromoproteins with a linear tetrapyrrole chromophore, similar to the ringed tetrapyrrole light-absorbing head group of chlorophyll. Phytochromes have two photo-interconvertible forms: Pr and Pfr. Pr absorbs red light (~667 nm) and is immediately converted to Pfr. Pfr absorbs far-red light (~730 nm) and is quickly converted back to Pr. Absorption of red or far-red light causes a massive change to the shape of the chromophore, altering the conformation and activity of the phytochrome protein to which it is bound. Pfr is the physiologically active form of the protein; therefore, exposure to red light yields physiological activity. Exposure to far-red light inhibits phytochrome activity. Together, the two forms represent the phytochrome system ([link]).

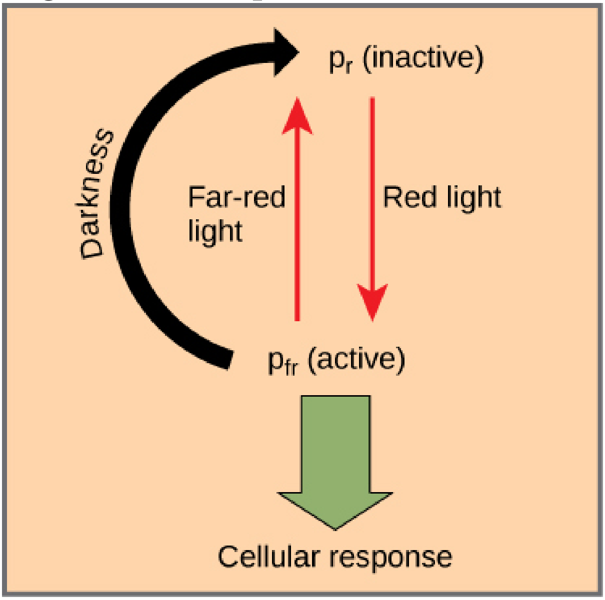
The phytochrome system acts as a biological light

switch. It monitors the level, intensity, duration, and color of environmental light. The effect of red light is reversible by immediately shining far-red light on the sample, which converts the chromoprotein to the inactive Pr form. Additionally, Pfr can slowly revert to Pr in the dark, or break down over time. In all instances, the physiological response induced by red light is reversed. The active form of phytochrome (Pfr) can directly activate other molecules in the cytoplasm, or it can be trafficked to the nucleus, where it directly activates or represses specific gene expression.

Once the phytochrome system evolved, plants adapted it to serve a variety of needs. Unfiltered, full sunlight contains much more red light than far-red light. Because chlorophyll absorbs strongly in the red region of the visible spectrum, but not in the far-red region, any plant in the shade of another plant on the forest floor will be exposed to red-depleted, far-red-enriched light. The preponderance of far-red light converts phytochrome in the shaded leaves to the Pr (inactive) form, slowing growth. The nearest non-shaded (or even less-shaded) areas on the forest floor have more red light; leaves exposed to these areas sense the red light, which activates the Pfr form and induces growth. In short, plant shoots use the phytochrome system to grow away from shade and towards light. Because competition for light is so fierce in a dense plant community, the evolutionary advantages of the

phytochrome system are obvious.

In seeds, the phytochrome system is not used to determine direction and quality of light (shaded versus unshaded). Instead, is it used merely to determine if there is any light at all. This is especially important in species with very small seeds, such as lettuce. Because of their size, lettuce seeds have few food reserves. Their seedlings cannot grow for long before they run out of fuel. If they germinated even a centimeter under the soil surface, the seedling would never make it into the sunlight and would die. In the dark, phytochrome is in the Pr (inactive form) and the seed will not germinate; it will only germinate if exposed to light at the surface of the soil. Upon exposure to light, Pr is converted to Pfr and germination proceeds.



Plants also use the phytochrome system to sense the change of season. Photoperiodism is a biological response to the timing and duration of day and night. It controls flowering, setting of winter buds, and vegetative growth. Detection of seasonal changes is crucial to plant survival. Although temperature and light intensity influence plant growth, they are not reliable indicators of season because they may vary from one year to the next. Day length is a better indicator of the time of year.

As stated above, unfiltered sunlight is rich in red light but deficient in far-red light. Therefore, at dawn, all the phytochrome molecules in a leaf quickly convert to the active Pfr form, and remain in that form until sunset. In the dark, the Pfr form takes hours to slowly revert back to the Pr form. If the night is long (as in winter), all of the Pfr form reverts. If the night is short (as in summer), a considerable amount of Pfr may remain at sunrise. By sensing the Pr/Pfr ratio at dawn, a plant can determine the length of the day/night cycle. In addition, leaves retain that information for several days, allowing a comparison between the length of the previous night and the preceding several nights. Shorter nights indicate springtime to the plant; when the nights become longer, autumn is approaching. This information, along with sensing temperature and water availability, allows plants to determine the time of the year and adjust their physiology accordingly. Short-day (long-night)

plants use this information to flower in the late summer and early fall, when nights exceed a critical length (often eight or fewer hours). Long-day (short-night) plants flower during the spring, when darkness is less than a critical length (often eight to 15 hours). Not all plants use the phytochrome system in this way. Flowering in day-neutral plants is not regulated by daylength.

Career Connection

Horticulturalist

The word “horticulturist” comes from the Latin words for garden (*hortus*) and culture (*cultura*). This career has been revolutionized by progress made in the understanding of plant responses to environmental stimuli. Growers of crops, fruit, vegetables, and flowers were previously constrained by having to time their sowing and harvesting according to the season. Now, horticulturists can manipulate plants to increase leaf, flower, or fruit production by understanding how environmental factors affect plant growth and development.

Greenhouse management is an essential component of a horticulturist’s education. To lengthen the night, plants are covered with a blackout shade cloth. Long-day plants are irradiated with red light in winter to promote early flowering. For example, fluorescent (cool white) light high in blue

wavelengths encourages leafy growth and is excellent for starting seedlings. Incandescent lamps (standard light bulbs) are rich in red light, and promote flowering in some plants. The timing of fruit ripening can be increased or delayed by applying plant hormones. Recently, considerable progress has been made in the development of plant breeds that are suited to different climates and resistant to pests and transportation damage. Both crop yield and quality have increased as a result of practical applications of the knowledge of plant responses to external stimuli and hormones. Horticulturists find employment in private and governmental laboratories, greenhouses, botanical gardens, and in the production or research fields. They improve crops by applying their knowledge of genetics and plant physiology. To prepare for a horticulture career, students take classes in botany, plant physiology, plant pathology, landscape design, and plant breeding. To complement these traditional courses, horticulture majors add studies in economics, business, computer science, and communications.

The Blue Light Responses

Phototropism—the directional bending of a plant toward or away from a light source—is a response to blue wavelengths of light. Positive phototropism

is growth towards a light source ([link]), while negative phototropism (also called skototropism) is growth away from light.

The aptly-named **phototropins** are protein-based receptors responsible for mediating the phototropic response. Like all plant photoreceptors, phototropins consist of a protein portion and a light-absorbing portion, called the chromophore. In phototropins, the chromophore is a covalently-bound molecule of flavin; hence, phototropins belong to a class of proteins called flavoproteins.

Other responses under the control of phototropins are leaf opening and closing, chloroplast movement, and the opening of stomata. However, of all responses controlled by phototropins, phototropism has been studied the longest and is the best understood.

In their 1880 treatise *The Power of Movements in Plants*, Charles Darwin and his son Francis first described phototropism as the bending of seedlings toward light. Darwin observed that light was perceived by the tip of the plant (the apical meristem), but that the response (bending) took place in a different part of the plant. They concluded that the signal had to travel from the apical meristem to the base of the plant.

In 1913, Peter Boysen-Jensen demonstrated that a chemical signal produced in the plant tip was responsible for the bending at the base. He cut off the tip of a seedling, covered the cut section with a layer of gelatin, and then replaced the tip. The seedling bent toward the light when illuminated. However, when impermeable mica flakes were inserted between the tip and the cut base, the seedling did not bend. A refinement of the experiment showed that the signal traveled on the shaded side of the seedling. When the mica plate was inserted on the illuminated side, the plant did bend towards the light. Therefore, the chemical signal was a growth stimulant because the phototropic response involved faster cell elongation on the shaded side than on the illuminated side. We now know that as light passes through a plant stem, it is diffracted and generates phototropin activation across the stem. Most activation occurs on the lit

side, causing the plant hormone indole acetic acid (IAA) to accumulate on the shaded side. Stem cells elongate under influence of IAA.

Cryptochromes are another class of blue-light absorbing photoreceptors that also contain a flavin-based chromophore. Cryptochromes set the plants' 24-hour activity cycle, also know as its circadian rhythem, using blue light cues. There is some evidence that cryptochromes work together with phototropins to mediate the phototropic response.

Link to Learning

Use the navigation menu in the left panel of this website to view images of plants in motion.

Plant Responses to Gravity

Whether or not they germinate in the light or in total darkness, shoots usually sprout up from the ground, and roots grow downward into the ground. A plant laid on its side in the dark will send shoots upward when given enough time. Gravitropism ensures that roots grow into the soil and that shoots grow toward sunlight. Growth of the shoot apical tip

upward is called **negative gravitropism**, whereas growth of the roots downward is called **positive gravitropism**.

Amyloplasts (also known as **statoliths**) are specialized plastids that contain starch granules and settle downward in response to gravity. Amyloplasts are found in shoots and in specialized cells of the root cap. When a plant is tilted, the statoliths drop to the new bottom cell wall. A few hours later, the shoot or root will show growth in the new vertical direction.

The mechanism that mediates gravitropism is reasonably well understood. When amyloplasts settle to the bottom of the gravity-sensing cells in the root or shoot, they physically contact the endoplasmic reticulum (ER), causing the release of calcium ions from inside the ER. This calcium signaling in the cells causes polar transport of the plant hormone IAA to the bottom of the cell. In roots, a high concentration of IAA inhibits cell elongation. The effect slows growth on the lower side of the root, while cells develop normally on the upper side. IAA has the opposite effect in shoots, where a higher concentration at the lower side of the shoot stimulates cell expansion, causing the shoot to grow up. After the shoot or root begin to grow vertically, the amyloplasts return to their normal position. Other hypotheses—involving the entire cell in the gravitropism effect—have been

proposed to explain why some mutants that lack amyloplasts may still exhibit a weak gravitropic response.

In grapes, application of gibberellic acid increases the size of fruit and loosens clustering. (credit: Bob Nichols, USDA)

Growth Responses

A plant’s sensory response to external stimuli relies on chemical messengers (hormones). Plant hormones affect all aspects of plant life, from flowering to fruit setting and maturation, and from phototropism to leaf fall. Potentially every cell in a plant can produce plant hormones. They can act in their cell of origin or be transported to other portions of the plant body, with many plant responses involving the synergistic or antagonistic interaction of two or more hormones. In contrast, animal hormones are produced in specific glands and transported to a distant site for action, and they act alone.

Plant hormones are a group of unrelated chemical substances that affect plant morphogenesis. Five major plant hormones are traditionally described: auxins (particularly IAA), cytokinins, gibberellins, ethylene, and abscisic acid. In addition, other nutrients and environmental conditions can be characterized as growth factors.

Auxins

The term auxin is derived from the Greek word *auxein*, which means "to grow." **Auxins** are the main hormones responsible for cell elongation in phototropism and gravitropism. They also control the differentiation of meristem into vascular tissue, and promote leaf development and arrangement. While many synthetic auxins are used as herbicides, IAA is the only naturally occurring auxin that shows physiological activity. Apical dominance—the inhibition of lateral bud formation—is triggered by auxins produced in the apical meristem. Flowering, fruit setting and ripening, and inhibition of **abscission** (leaf falling) are other plant responses under the direct or indirect control of auxins. Auxins also act as a relay for the effects of the blue light and red/far-red responses.

Commercial use of auxins is widespread in plant nurseries and for crop production. IAA is used as a rooting hormone to promote growth of adventitious roots on cuttings and detached leaves. Applying synthetic auxins to tomato plants in greenhouses promotes normal fruit development. Outdoor application of auxin promotes synchronization of fruit setting and dropping to coordinate the harvesting season. Fruits such as seedless cucumbers can be induced to set fruit by treating unfertilized plant flowers with auxins.

Cytokinins

The effect of cytokinins was first reported when it was found that adding the liquid endosperm of coconuts to developing plant embryos in culture stimulated their growth. The stimulating growth factor was found to be **cytokinin**, a hormone that promotes cytokinesis (cell division). Almost 200 naturally occurring or synthetic cytokinins are known to date. Cytokinins are most abundant in growing tissues, such as roots, embryos, and fruits, where cell division is occurring. Cytokinins are known to delay senescence in leaf tissues, promote mitosis, and stimulate differentiation of the meristem in shoots and roots. Many effects on plant development are under the influence of cytokinins, either in conjunction with auxin or another hormone. For example, apical dominance seems to result from a balance between auxins that inhibit lateral buds, and cytokinins that promote bushier growth.

Gibberellins

Gibberellins (GAs) are a group of about 125 closely related plant hormones that stimulate shoot elongation, seed germination, and fruit and flower maturation. GAs are synthesized in the root and stem apical meristems, young leaves, and seed embryos. In urban areas, GA antagonists are sometimes applied to trees under power lines to

control growth and reduce the frequency of pruning.

GAs break dormancy (a state of inhibited growth and development) in the seeds of plants that require exposure to cold or light to germinate. Abscisic acid is a strong antagonist of GA action. Other effects of GAs include gender expression, seedless fruit development, and the delay of senescence in leaves and fruit. Seedless grapes are obtained through standard breeding methods and contain inconspicuous seeds that fail to develop. Because GAs are produced by the seeds, and because fruit development and stem elongation are under GA control, these varieties of grapes would normally produce small fruit in compact clusters. Maturing grapes are routinely treated with GA to promote larger fruit size, as well as looser bunches (longer stems), which reduces the instance of mildew infection ([link]).

Abscisic Acid

The plant hormone **abscisic acid** (ABA) was first discovered as the agent that causes the abscission or dropping of cotton bolls. However, more recent studies indicate that ABA plays only a minor role in the abscission process. ABA accumulates as a

response to stressful environmental conditions, such as dehydration, cold temperatures, or shortened day lengths. Its activity counters many of the growth-promoting effects of GAs and auxins. ABA inhibits stem elongation and induces dormancy in lateral buds.

ABA induces dormancy in seeds by blocking germination and promoting the synthesis of storage proteins. Plants adapted to temperate climates require a long period of cold temperature before seeds germinate. This mechanism protects young plants from sprouting too early during unseasonably warm weather in winter. As the hormone gradually breaks down over winter, the seed is released from dormancy and germinates when conditions are favorable in spring. Another effect of ABA is to promote the development of winter buds; it mediates the conversion of the apical meristem into a dormant bud. Low soil moisture causes an increase in ABA, which causes stomata to close, reducing water loss in winter buds.

Ethylene

Ethylene is associated with fruit ripening, flower wilting, and leaf fall. Ethylene is unusual because it is a volatile gas (C_2H_4). Hundreds of years ago, when gas street lamps were installed in city streets, trees that grew close to lamp posts developed twisted, thickened trunks and shed their leaves

earlier than expected. These effects were caused by ethylene volatilizing from the lamps.

Aging tissues (especially senescing leaves) and nodes of stems produce ethylene. The best-known effect of the hormone, however, is the promotion of fruit ripening. Ethylene stimulates the conversion of starch and acids to sugars. Some people store unripe fruit, such as avocadoes, in a sealed paper bag to accelerate ripening; the gas released by the first fruit to mature will speed up the maturation of the remaining fruit. Ethylene also triggers leaf and fruit abscission, flower fading and dropping, and promotes germination in some cereals and sprouting of bulbs and potatoes.

Ethylene is widely used in agriculture. Commercial fruit growers control the timing of fruit ripening with application of the gas. Horticulturalists inhibit leaf dropping in ornamental plants by removing ethylene from greenhouses using fans and ventilation.

Nontraditional Hormones

Recent research has discovered a number of compounds that also influence plant development. Their roles are less understood than the effects of the major hormones described so far.

Jasmonates play a major role in defense responses

to herbivory. Their levels increase when a plant is wounded by a predator, resulting in an increase in toxic secondary metabolites. They contribute to the production of volatile compounds that attract natural enemies of predators. For example, chewing of tomato plants by caterpillars leads to an increase in jasmonic acid levels, which in turn triggers the release of volatile compounds that attract predators of the pest.

Oligosaccharins also play a role in plant defense against bacterial and fungal infections. They act locally at the site of injury, and can also be transported to other tissues. **Strigolactones** promote seed germination in some species and inhibit lateral apical development in the absence of auxins. Strigolactones also play a role in the establishment of mycorrhizae, a mutualistic association of plant roots and fungi. Brassinosteroids are important to many developmental and physiological processes. Signals between these compounds and other hormones, notably auxin and GAs, amplifies their physiological effect. Apical dominance, seed germination, gravitropism, and resistance to freezing are all positively influenced by hormones. Root growth and fruit dropping are inhibited by steroids.

Plant Responses to Wind and Touch

The shoot of a pea plant winds around a trellis, while a tree grows on an angle in response to strong prevailing winds. These are examples of how plants respond to touch or wind.

The movement of a plant subjected to constant directional pressure is called **thigmotropism**, from the Greek words *thigma* meaning “touch,” and *tropism* implying “direction.” Tendrils are one example of this. The meristematic region of tendrils is very touch sensitive; light touch will evoke a quick coiling response. Cells in contact with a support surface contract, whereas cells on the opposite side of the support expand ([link]). Application of jasmonic acid is sufficient to trigger tendril coiling without a mechanical stimulus.

A **thigmonastic** response is a touch response independent of the direction of stimulus [link]. In the Venus flytrap, two modified leaves are joined at a hinge and lined with thin fork-like tines along the outer edges. Tiny hairs are located inside the trap. When an insect brushes against these trigger hairs, touching two or more of them in succession, the leaves close quickly, trapping the prey. Glands on the leaf surface secrete enzymes that slowly digest the insect. The released nutrients are absorbed by the leaves, which reopen for the next meal.

Thigmomorphogenesis is a slow developmental change in the shape of a plant subjected to

continuous mechanical stress. When trees bend in the wind, for example, growth is usually stunted and the trunk thickens. Strengthening tissue, especially xylem, is produced to add stiffness to resist the wind's force. Researchers hypothesize that mechanical strain induces growth and differentiation to strengthen the tissues. Ethylene and jasmonate are likely involved in thigmomorphogenesis.

> Link to Learning
>
> Use the menu at the left to navigate to three short movies: a Venus fly trap capturing prey, the progressive closing of sensitive plant leaflets, and the twining of tendrils.

Defense Responses against Herbivores and Pathogens

Plants face two types of enemies: herbivores and pathogens. Herbivores both large and small use plants as food, and actively chew them. Pathogens are agents of disease. These infectious microorganisms, such as fungi, bacteria, and

nematodes, live off of the plant and damage its tissues. Plants have developed a variety of strategies to discourage or kill attackers.

The first line of defense in plants is an intact and impenetrable barrier. Bark and the waxy cuticle can protect against predators. Other adaptations against herbivory include thorns, which are modified branches, and spines, which are modified leaves. They discourage animals by causing physical damage and inducing rashes and allergic reactions. A plant's exterior protection can be compromised by mechanical damage, which may provide an entry point for pathogens. If the first line of defense is breached, the plant must resort to a different set of defense mechanisms, such as toxins and enzymes.

Secondary metabolites are compounds that are not directly derived from photosynthesis and are not necessary for respiration or plant growth and development. Many metabolites are toxic, and can even be lethal to animals that ingest them. Some metabolites are alkaloids, which discourage predators with noxious odors (such as the volatile oils of mint and sage) or repellent tastes (like the bitterness of quinine). Other alkaloids affect herbivores by causing either excessive stimulation (caffeine is one example) or the lethargy associated with opioids. Some compounds become toxic after ingestion. For instance, glycol cyanide in the cassava root releases cyanide only upon ingestion;

the nearly 500 million humans who rely on cassava for nutrition must be certain to process the root properly before eating.

Mechanical wounding and predator attacks activate defense and protection mechanisms both in the damaged tissue and at sites farther from the injury location. Some defense reactions occur within minutes: others over several hours. The infected and surrounding cells may die, thereby stopping the spread of infection.

Long-distance signaling elicits a systemic response aimed at deterring the predator. As tissue is damaged, jasmonates may promote the synthesis of compounds that are toxic to predators. Jasmonates also elicit the synthesis of volatile compounds that attract parasitoids, which are insects that spend their developing stages in or on another insect, and eventually kill their host. The plant may activate abscission of injured tissue if it is damaged beyond repair.

Section Summary

Plants respond to light by changes in morphology and activity. Irradiation by red light converts the photoreceptor phytochrome to its far-red light-absorbing form—Pfr. This form controls germination and flowering in response to length of day, as well

as triggers photosynthesis in dormant plants or those that just emerged from the soil. Blue-light receptors, cryptochromes, and phototropins are responsible for phototropism. Amyloplasts, which contain heavy starch granules, sense gravity. Shoots exhibit negative gravitropism, whereas roots exhibit positive gravitropism. Plant hormones—naturally occurring compounds synthesized in small amounts —can act both in the cells that produce them and in distant tissues and organs. Auxins are responsible for apical dominance, root growth, directional growth toward light, and many other growth responses. Cytokinins stimulate cell division and counter apical dominance in shoots. Gibberellins inhibit dormancy of seeds and promote stem growth. Abscisic acid induces dormancy in seeds and buds, and protects plants from excessive water loss by promoting stomatal closure. Ethylene gas speeds up fruit ripening and dropping of leaves. Plants respond to touch by rapid movements (thigmotropy and thigmonasty) and slow differential growth (thigmomorphogenesis). Plants have evolved defense mechanisms against predators and pathogens. Physical barriers like bark and spines protect tender tissues. Plants also have chemical defenses, including toxic secondary metabolites and hormones, which elicit additional defense mechanisms.

Review Questions

The main photoreceptor that triggers phototropism is a _____.

1. phytochrome
2. cryptochrome
3. phototropin
4. carotenoid

C

Phytochrome is a plant pigment protein that:

1. mediates plant infection
2. promotes plant growth
3. mediates morphological changes in response to red and far-red light
4. inhibits plant growth

C

A mutant plant has roots that grow in all directions. Which of the following organelles would you expect to be missing in the cell?

1. mitochondria
2. amyloplast
3. chloroplast

4. nucleus

B

After buying green bananas or unripe avocadoes, they can be kept in a brown bag to ripen. The hormone released by the fruit and trapped in the bag is probably:

1. abscisic acid
2. cytokinin
3. ethylene
4. gibberellic acid

C

A decrease in the level of which hormone releases seeds from dormancy?

1. abscisic acid
2. cytokinin
3. ethylene
4. gibberellic acid

A

A seedling germinating under a stone grows at an angle away from the stone and upward. This response to touch is called _____.

1. gravitropism
2. thigmonasty
3. thigmotropism
4. skototropism

C

Critical Thinking Questions

Owners and managers of plant nurseries have to plan lighting schedules for a long-day plant that will flower in February. What lighting periods will be most effective? What color of light should be chosen?

A long-day plant needs a higher proportion of the Pfr form to Pr form of phytochrome. The plant requires long periods of illumination with light enriched in the red range of the spectrum.

What are the major benefits of gravitropism for

a germinating seedling?

Gravitropism will allow roots to dig deep into the soil to find water and minerals, whereas the seedling will grow towards light to enable photosynthesis.

Fruit and vegetable storage facilities are usually refrigerated and well ventilated. Why are these conditions advantageous?

Refrigeration slows chemical reactions, including fruit maturation. Ventilation removes the ethylene gas that speeds up fruit ripening.

Stomata close in response to bacterial infection. Why is this response a mechanism of defense for the plant? Which hormone is most likely to mediate this response?

To prevent further entry of pathogens, stomata close, even if they restrict entry of CO_2. Some pathogens secrete virulence factors that inhibit the closing of stomata. Abscisic acid is the stress hormone responsible for inducing closing of stomata.

Glossary

abscisic acid (ABA)
: plant hormone that induces dormancy in seeds and other organs

abscission
: physiological process that leads to the fall of a plant organ (such as leaf or petal drop)

auxin
: plant hormone that influences cell elongation (in phototropism), gravitropism, apical dominance, and root growth

chromophore
: molecule that absorbs light

cryptochrome
: protein that absorbs light in the blue and ultraviolet regions of the light spectrum

cytokinin
: plant hormone that promotes cell division

ethylene
: volatile plant hormone that is associated with fruit ripening, flower wilting, and leaf fall

gibberellin (GA)
: plant hormone that stimulates shoot elongation, seed germination, and the

maturation and dropping of fruit and flowers

jasmonates
: small family of compounds derived from the fatty acid linoleic acid

negative gravitropism
: growth away from Earth's gravity

oligosaccharin
: hormone important in plant defenses against bacterial and fungal infections

photomorphogenesis
: growth and development of plants in response to light

photoperiodism
: occurrence of plant processes, such as germination and flowering, according to the time of year

phototropin
: blue-light receptor that promotes phototropism, stomatal opening and closing, and other responses that promote photosynthesis

phototropism
: directional bending of a plant toward a light source

phytochrome

plant pigment protein that exists in two reversible forms (Pr and Pfr) and mediates morphologic changes in response to red light

positive gravitropism
: growth toward Earth’s gravitational center

statolith
: (also, **amyloplast**) plant organelle that contains heavy starch granules

strigolactone
: hormone that promotes seed germination in some species and inhibits lateral apical development in the absence of auxins

thigmomorphogenesis
: developmental response to touch

thigmonastic
: directional growth of a plant independent of the direction in which contact is applied

thigmotropism
: directional growth of a plant in response to constant contact

Introduction

class="introduction" For this (a) squash seedling (*Cucurbita maxima*) to develop into a mature plant bearing its (b) fruit, numerous nutritional requirements must be met. (credit a: modification of work by Julian Colton; credit b: modification of work by "Wildfeuer"/Wikimedia Commons)

(a) (b)

Cucurbitaceae is a family of plants first cultivated in Mesoamerica, although several species are native to North America. The family includes many edible species, such as squash and pumpkin, as well as inedible gourds. In order to grow and develop into mature, fruit-bearing plants, many requirements must be met and events must be coordinated. Seeds must germinate under the right conditions in the soil; therefore, temperature, moisture, and soil quality are important factors that play a role in germination and seedling development. Soil quality and climate are significant to plant distribution and growth. The young seedling will eventually grow into a mature plant, and the roots will absorb nutrients and water from the soil. At the same time, the aboveground parts of the plant will absorb

carbon dioxide from the atmosphere and use energy from sunlight to produce organic compounds through photosynthesis. This chapter will explore the complex dynamics between plants and soils, and the adaptations that plants have evolved to make better use of nutritional resources.

Nutritional Requirements of Plants

By the end of this section, you will be able to do the following:

- Describe how plants obtain nutrients
- List the elements and compounds required for proper plant nutrition
- Describe an essential nutrient

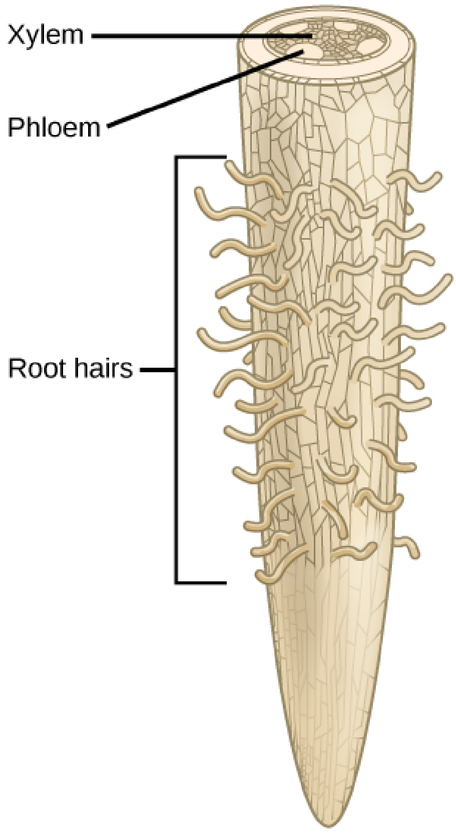
Plants are unique organisms that can absorb nutrients and water through their root system, as well as carbon dioxide from the atmosphere. Soil quality and climate are the major determinants of plant distribution and growth. The combination of soil nutrients, water, and carbon dioxide, along with sunlight, allows plants to grow.

Water is absorbed through the root hairs and moves up the xylem to the leaves.

The Chemical Composition of Plants

Since plants require nutrients in the form of elements such as carbon and potassium, it is important to understand the chemical composition of plants. The majority of volume in a plant cell is water; it typically comprises 80 to 90 percent of the plant's total weight. Soil is the water source for land plants, and can be an abundant source of water, even if it appears dry. Plant roots absorb water from the soil through root hairs and transport it up to the

leaves through the xylem. As water vapor is lost from the leaves, the process of transpiration and the polarity of water molecules (which enables them to form hydrogen bonds) draws more water from the roots up through the plant to the leaves ([link]). Plants need water to support cell structure, for metabolic functions, to carry nutrients, and for photosynthesis.



Plant cells need essential substances, collectively called nutrients, to sustain life. Plant nutrients may be composed of either organic or inorganic compounds. An **organic compound** is a chemical

compound that contains carbon, such as carbon dioxide obtained from the atmosphere. Carbon that was obtained from atmospheric CO_2 composes the majority of the dry mass within most plants. An **inorganic compound** does not contain carbon and is not part of, or produced by, a living organism. Inorganic substances, which form the majority of the soil solution, are commonly called minerals: those required by plants include nitrogen (N) and potassium (K) for structure and regulation.

Cellulose, the main structural component of the plant cell wall, makes up over thirty percent of plant matter. It is the most abundant organic compound on earth. Nutrient deficiency is evident in the symptoms these plants show. This (a) grape tomato suffers from blossom end rot caused by calcium deficiency. The yellowing in this (b) *Frangula alnus* results from magnesium deficiency. Inadequate magnesium also leads to (c) intervenal chlorosis, seen here in a sweetgum leaf. This (d) palm is affected by potassium deficiency. (credit c: modification of work by Jim Conrad; credit d: modification of work by Malcolm Manners)

Essential Nutrients

Plants require only light, water, and about 20 elements to support all their biochemical needs: these 20 elements are called essential nutrients ([link]). For an element to be regarded as **essential**, three criteria are required: 1) a plant cannot

complete its life cycle without the element; 2) no other element can perform the function of the element; and 3) the element is directly involved in plant nutrition.

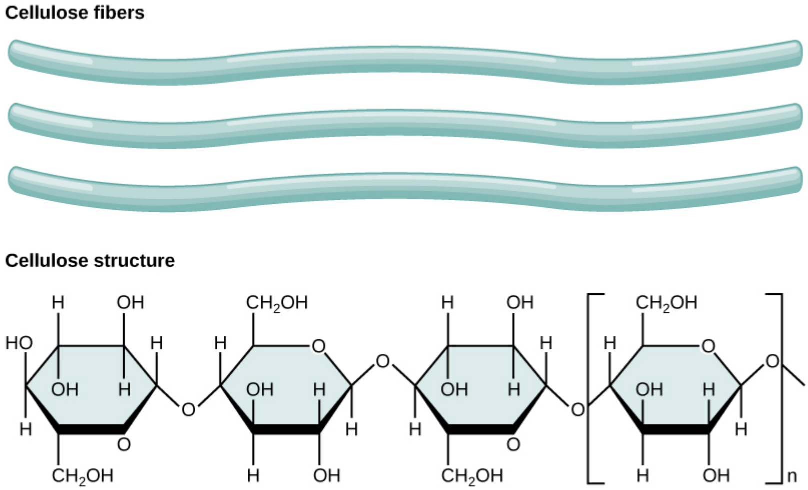
Essential Elements for Plant Growth

Macronutrients	Micronutrients
Carbon (C)	Iron (Fe)
Hydrogen (H)	Manganese (Mn)
Oxygen (O)	Boron (B)
Nitrogen (N)	Molybdenum (Mo)
Phosphorus (P)	Copper (Cu)
Potassium (K)	Zinc (Zn)
Calcium (Ca)	Chlorine (Cl)
Magnesium (Mg)	Nickel (Ni)
Sulfur (S)	Cobalt (Co)
	Sodium (Na)
	Silicon (Si)

Macronutrients and Micronutrients

The essential elements can be divided into two groups: macronutrients and micronutrients. Nutrients that plants require in larger amounts are called **macronutrients**. About half of the essential

elements are considered macronutrients: carbon, hydrogen, oxygen, nitrogen, phosphorus, potassium, calcium, magnesium and sulfur. The first of these macronutrients, carbon (C), is required to form carbohydrates, proteins, nucleic acids, and many other compounds; it is therefore present in all macromolecules. On average, the dry weight (excluding water) of a cell is 50 percent carbon. As shown in [link], carbon is a key part of plant biomolecules.



The next most abundant element in plant cells is nitrogen (N); it is part of proteins and nucleic acids. Nitrogen is also used in the synthesis of some vitamins. Hydrogen and oxygen are macronutrients that are part of many organic compounds, and also form water. Oxygen is necessary for cellular respiration; plants use oxygen to store energy in the form of ATP. Phosphorus (P), another macromolecule, is necessary to synthesize nucleic

acids and phospholipids. As part of ATP, phosphorus enables food energy to be converted into chemical energy through oxidative phosphorylation. Likewise, light energy is converted into chemical energy during photophosphorylation in photosynthesis, and into chemical energy to be extracted during respiration. Sulfur is part of certain amino acids, such as cysteine and methionine, and is present in several coenzymes. Sulfur also plays a role in photosynthesis as part of the electron transport chain, where hydrogen gradients play a key role in the conversion of light energy into ATP. Potassium (K) is important because of its role in regulating stomatal opening and closing. As the openings for gas exchange, stomata help maintain a healthy water balance; a potassium ion pump supports this process.

Magnesium (Mg) and calcium (Ca) are also important macronutrients. The role of calcium is twofold: to regulate nutrient transport, and to support many enzyme functions. Magnesium is important to the photosynthetic process. These minerals, along with the micronutrients, which are described below, also contribute to the plant's ionic balance.

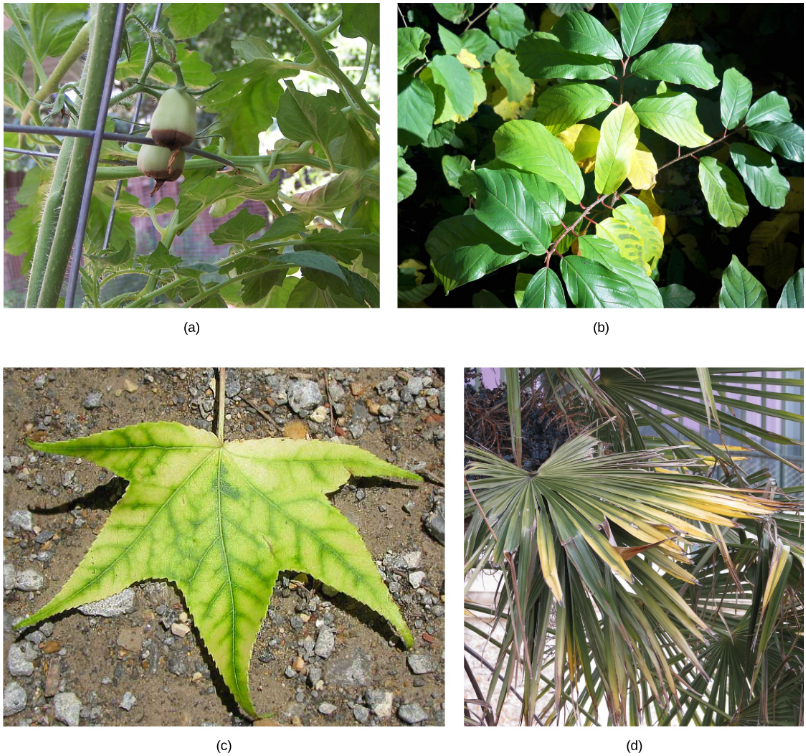
In addition to macronutrients, organisms require various elements in small amounts. These **micronutrients**, or trace elements, are present in very small quantities. They include boron (B),

chlorine (Cl), manganese (Mn), iron (Fe), zinc (Zn), copper (Cu), molybdenum (Mo), nickel (Ni), silicon (Si), and sodium (Na).

Deficiencies in any of these nutrients—particularly the macronutrients—can adversely affect plant growth ([link]). Depending on the specific nutrient, a lack can cause stunted growth, slow growth, or chlorosis (yellowing of the leaves). Extreme deficiencies may result in leaves showing signs of cell death.

Link to Learning

Visit this website to participate in an interactive experiment on plant nutrient deficiencies. You can adjust the amounts of N, P, K, Ca, Mg, and Fe that plants receive . . . and see what happens.



Everyday Connection

Plant physiologist Ray Wheeler checks onions being grown using hydroponic techniques. The other plants are Bibb lettuce (left) and radishes (right). Credit: NASA

Hydroponics

Hydroponics is a method of growing plants in a water-nutrient solution instead of soil. Since its advent, hydroponics has developed into a growing process that researchers often use. Scientists who are interested in studying plant nutrient deficiencies can use hydroponics to study the effects of different nutrient combinations under strictly controlled conditions. Hydroponics has also developed as a way to grow flowers, vegetables, and other crops in greenhouse environments. You might find hydroponically grown produce at your local grocery store. Today, many lettuces and tomatoes in your market have been hydroponically grown.

Section Summary

Plants can absorb inorganic nutrients and water through their root system, and carbon dioxide from the environment. The combination of organic compounds, along with water, carbon dioxide, and sunlight, produce the energy that allows plants to grow. Inorganic compounds form the majority of the soil solution. Plants access water though the soil. Water is absorbed by the plant root, transports nutrients throughout the plant, and maintains the structure of the plant. Essential elements are indispensable elements for plant growth. They are divided into macronutrients and micronutrients. The macronutrients plants require are carbon, nitrogen, hydrogen, oxygen, phosphorus, potassium, calcium, magnesium, and sulfur. Important micronutrients include iron, manganese, boron, molybdenum, copper, zinc, chlorine, nickel, cobalt, silicon, and sodium.

Review Questions

For an element to be regarded as essential, all of the following criteria must be met, except:

1. No other element can perform the function.

2. The element is directly involved in plant nutrition.
3. The element is inorganic.
4. The plant cannot complete its lifecycle without the element.

C

The nutrient that is part of carbohydrates, proteins, and nucleic acids, and that forms biomolecules, is ______.

1. nitrogen
2. carbon
3. magnesium
4. iron

B

Most ______ are necessary for enzyme function.

1. micronutrients
2. macronutrients
3. biomolecules
4. essential nutrients

A

What is the main water source for land plants?

1. rain
2. soil
3. biomolecules
4. essential nutrients

B

Critical Thinking Questions

What type of plant problems result from nitrogen and calcium deficiencies?

Deficiencies in these nutrients could result in stunted growth, slow growth, and chlorosis.

Research the life of Jan Babtista van Helmont. What did the van Helmont experiment show?

---

van Helmont showed that plants do not consume soil, which is correct. He also thought that plant growth and increased weight resulted from the intake of water, a conclusion that has since been disproven.

List two essential macronutrients and two essential micro nutrients.

Answers may vary. Essential macronutrients include carbon, hydrogen, oxygen, nitrogen, phosphorus, potassium, calcium, magnesium, and sulfur. Essential micronutrients include iron, manganese, boron, molybdenum, copper, zinc, chlorine, nickel, cobalt, sodium, and silicon.

Glossary

inorganic compound
: chemical compound that does not contain carbon; it is not part of or produced by a living organism

macronutrient
: nutrient that is required in large amounts for plant growth; carbon, hydrogen, oxygen, nitrogen, phosphorus, potassium, calcium, magnesium, and sulfur

micronutrient
: nutrient required in small amounts; also called trace element

organic compound

chemical compound that contains carbon

The Soil

By the end of this section, you will be able to do the following:

- Describe how soils are formed
- Explain soil composition
- Describe a soil profile

Plants obtain inorganic elements from the soil, which serves as a natural medium for land plants. **Soil** is the outer loose layer that covers the surface of Earth. Soil quality is a major determinant, along with climate, of plant distribution and growth. Soil quality depends not only on the chemical composition of the soil, but also the topography (regional surface features) and the presence of living organisms. In agriculture, the history of the soil, such as the cultivating practices and previous crops, modify the characteristics and fertility of that soil.

Soil develops very slowly over long periods of time, and its formation results from natural and environmental forces acting on mineral, rock, and organic compounds. Soils can be divided into two groups: **organic soils** are those that are formed from sedimentation and primarily composed of organic matter, while those that are formed from the weathering of rocks and are primarily composed of inorganic material are called **mineral soils**. Mineral soils are predominant in terrestrial ecosystems, where soils may be covered by water

for part of the year or exposed to the atmosphere.

Soil Composition

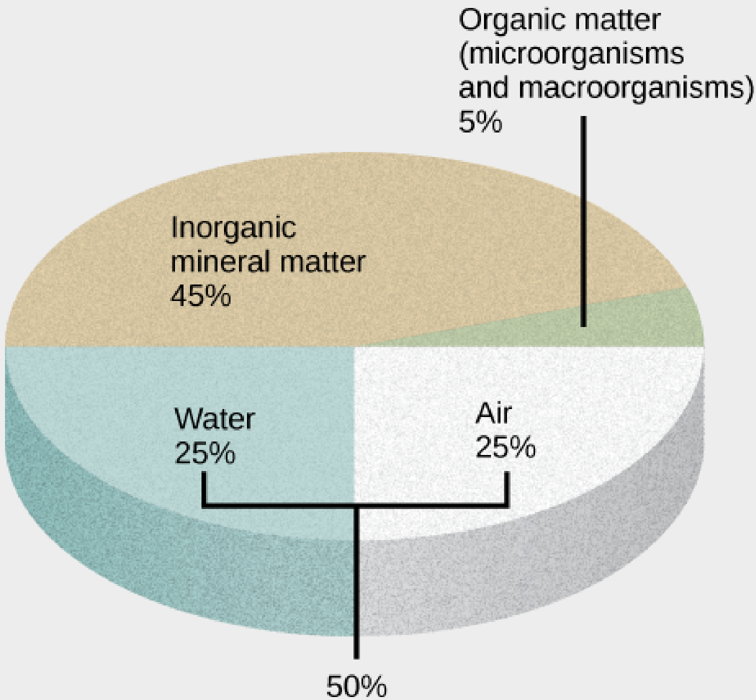
Soil consists of these major components ([link]):

- inorganic mineral matter, about 40 to 45 percent of the soil volume
- organic matter, about 5 percent of the soil volume
- water and air, about 50 percent of the soil volume

The amount of each of the four major components of soil depends on the amount of vegetation, soil compaction, and water present in the soil. A good healthy soil has sufficient air, water, minerals, and organic material to promote and sustain plant life.

Visual Connection

The four major components of soil are shown: inorganic minerals, organic matter, water, and air.



Soil compaction can result when soil is compressed by heavy machinery or even foot traffic. How might this compaction change the soil composition?

The organic material of soil, called **humus**, is made up of microorganisms (dead and alive), and dead animals and plants in varying stages of decay. Humus improves soil structure and provides plants with water and minerals. The inorganic material of soil consists of rock, slowly broken down into smaller particles that vary in size. Soil particles that are 0.1 to 2 mm in diameter are **sand**. Soil particles between 0.002 and 0.1 mm are called **silt**, and even

smaller particles, less than 0.002 mm in diameter, are called **clay**. Some soils have no dominant particle size and contain a mixture of sand, silt, and humus; these soils are called **loams**.

Link to Learning

Explore this interactive map from the USDA's National Cooperative Soil Survey to access soil data for almost any region in the United States.

Soil Formation

Soil formation is the consequence of a combination of biological, physical, and chemical processes. Soil should ideally contain 50 percent solid material and 50 percent pore space. About one-half of the pore space should contain water, and the other half should contain air. The organic component of soil serves as a cementing agent, returns nutrients to the plant, allows soil to store moisture, makes soil tillable for farming, and provides energy for soil microorganisms. Most soil microorganisms—bacteria, algae, or fungi—are dormant in dry soil, but become active once moisture is available.

Soil distribution is not homogenous because its formation results in the production of layers; together, the vertical section of a soil is called the **soil profile**. Within the soil profile, soil scientists define zones called horizons. A **horizon** is a soil layer with distinct physical and chemical properties that differ from those of other layers. Five factors account for soil formation: parent material, climate, topography, biological factors, and time.

Parent Material

The organic and inorganic material in which soils form is the **parent material**. Mineral soils form directly from the weathering of **bedrock**, the solid rock that lies beneath the soil, and therefore, they have a similar composition to the original rock. Other soils form in materials that came from elsewhere, such as sand and glacial drift. Materials located in the depth of the soil are relatively unchanged compared with the deposited material. Sediments in rivers may have different characteristics, depending on whether the stream moves quickly or slowly. A fast-moving river could have sediments of rocks and sand, whereas a slow-moving river could have fine-textured material, such as clay.

Climate

Temperature, moisture, and wind cause different

patterns of weathering and therefore affect soil characteristics. The presence of moisture and nutrients from weathering will also promote biological activity: a key component of a quality soil.

Topography

Regional surface features (familiarly called “the lay of the land”) can have a major influence on the characteristics and fertility of a soil. Topography affects water runoff, which strips away parent material and affects plant growth. Steeps soils are more prone to erosion and may be thinner than soils that are relatively flat or level.

Biological factors

The presence of living organisms greatly affects soil formation and structure. Animals and microorganisms can produce pores and crevices, and plant roots can penetrate into crevices to produce more fragmentation. Plant secretions promote the development of microorganisms around the root, in an area known as the **rhizosphere**. Additionally, leaves and other material that fall from plants decompose and contribute to soil composition.

Time

Time is an important factor in soil formation because soils develop over long periods. Soil formation is a dynamic process. Materials are deposited over time, decompose, and transform into other materials that can be used by living organisms or deposited onto the surface of the soil.

The San Joaquin soil profile has an O horizon, A horizon, B horizon, and C horizon. (credit: modification of work by USDA)

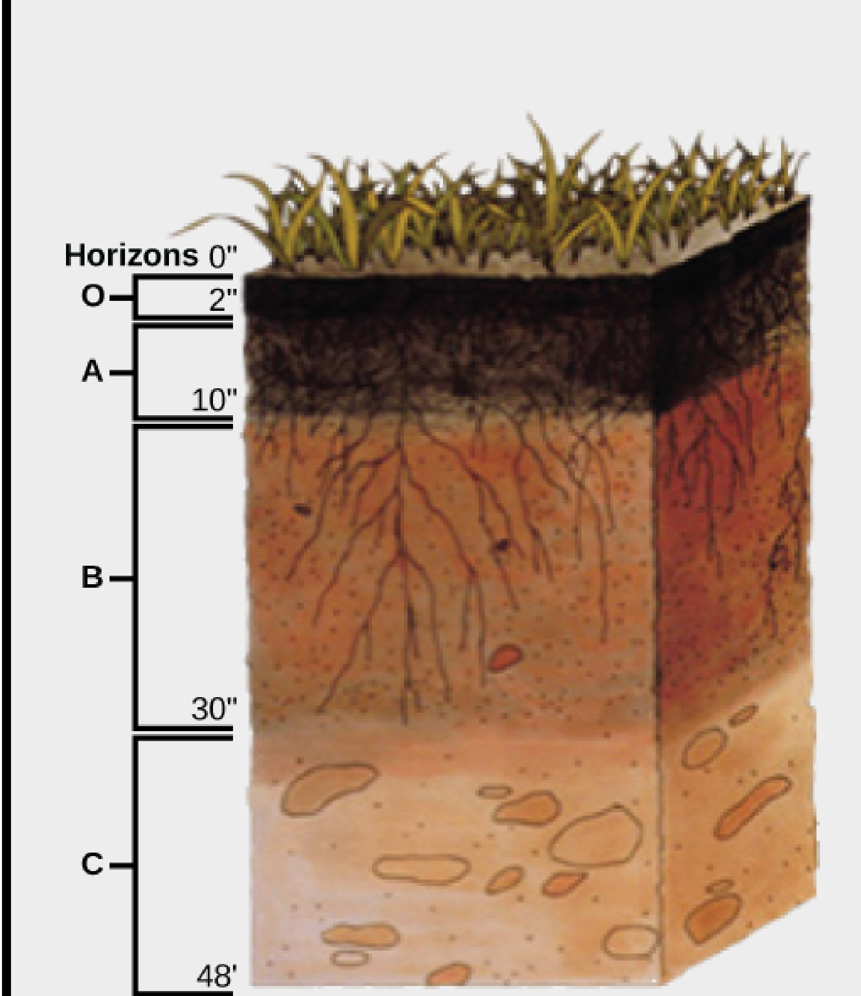
Physical Properties of the Soil

Soils are named and classified based on their horizons. The soil profile has four distinct layers: 1) O horizon; 2) A horizon; 3) B horizon, or subsoil; and 4) C horizon, or soil base ([link]). The **O horizon** has freshly decomposing organic matter—humus—at its surface, with decomposed vegetation at its base. Humus enriches the soil with nutrients and enhances soil moisture retention. Topsoil—the top layer of soil—is usually two to three inches deep, but this depth can vary considerably. For instance, river deltas like the Mississippi River delta have deep layers of topsoil. Topsoil is rich in organic material; microbial processes occur there, and it is the "workhorse" of plant production. The **A horizon** consists of a mixture of organic material with inorganic products of weathering, and it is therefore the beginning of true mineral soil. This horizon is typically darkly colored because of the presence of organic matter. In this area, rainwater

percolates through the soil and carries materials from the surface. The **B horizon** is an accumulation of mostly fine material that has moved downward, resulting in a dense layer in the soil. In some soils, the B horizon contains nodules or a layer of calcium carbonate. The **C horizon**, or soil base, includes the parent material, plus the organic and inorganic material that is broken down to form soil. The parent material may be either created in its natural place, or transported from elsewhere to its present location. Beneath the C horizon lies bedrock.

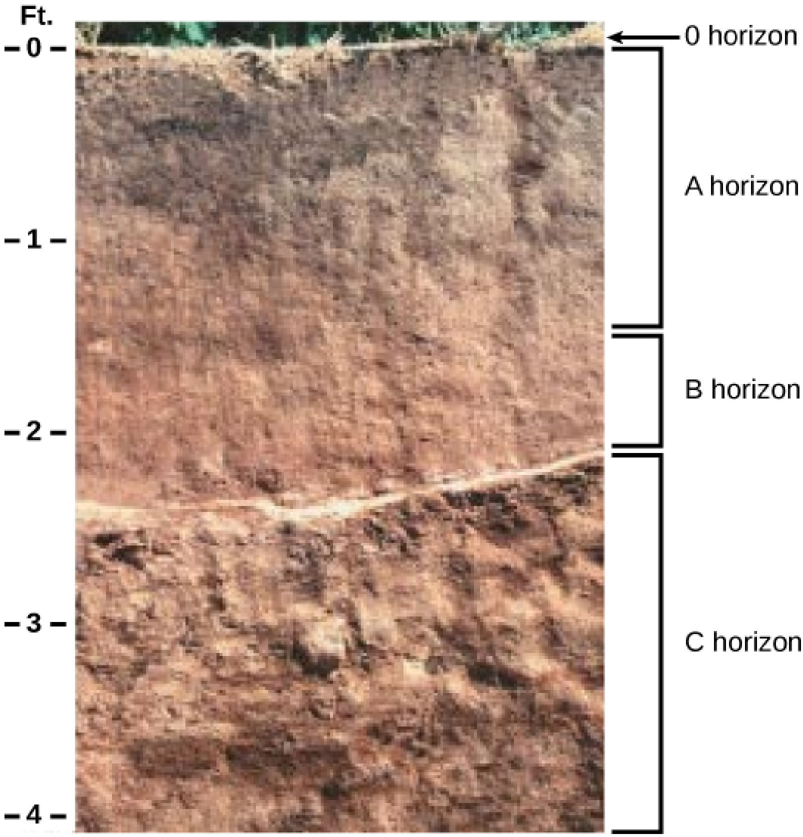
Visual Connection

This soil profile shows the different soil layers (O horizon, A horizon, B horizon, and C horizon) found in typical soils. (credit: modification of work by USDA)



Which horizon is considered the topsoil, and which is considered the subsoil?

Some soils may have additional layers, or lack one of these layers. The thickness of the layers is also variable, and depends on the factors that influence soil formation. In general, immature soils may have O, A, and C horizons, whereas mature soils may display all of these, plus additional layers ([link]).



Career Connections

Soil Scientist

A soil scientist studies the biological components, physical and chemical properties, distribution, formation, and morphology of soils. Soil scientists need to have a strong background in physical and life sciences, plus a foundation in mathematics. They may work for federal or state agencies, academia, or the private sector. Their work may involve collecting data, carrying out research,

interpreting results, inspecting soils, conducting soil surveys, and recommending soil management programs.

This soil scientist is studying the horizons and composition of soil at a research site. (credit: USDA)

Many soil scientists work both in an office and in the field. According to the United States Department of Agriculture (USDA): “a soil scientist needs good observation skills to analyze and

determine the characteristics of different types of soils. Soil types are complex and the geographical areas a soil scientist may survey are varied. Aerial photos or various satellite images are often used to research the areas. Computer skills and geographic information systems (GIS) help the scientist to analyze the multiple facets of geomorphology, topography, vegetation, and climate to discover the patterns left on the landscape."[footnote] Soil scientists play a key role in understanding the soil's past, analyzing present conditions, and making recommendations for future soil-related practices.
National Resources Conservation Service / United States Department of Agriculture. "Careers in Soil Science." http://openstax.org/l/NRCS

Section Summary

Plants obtain mineral nutrients from the soil. Soil is the outer loose layer that covers the surface of Earth. Soil quality depends on the chemical composition of the soil, the topography, the presence of living organisms, the climate, and time. Agricultural practice and history may also modify the characteristics and fertility of soil. Soil consists of four major components: 1) inorganic mineral matter, 2) organic matter, 3) water and air, and 4)

living matter. The organic material of soil is made of humus, which improves soil structure and provides water and minerals. Soil inorganic material consists of rock slowly broken down into smaller particles that vary in size, such as sand, silt, and loam.

Soil formation results from a combination of biological, physical, and chemical processes. Soil is not homogenous because its formation results in the production of layers called a soil profile. Factors that affect soil formation include: parent material, climate, topography, biological factors, and time. Soils are classified based on their horizons, soil particle size, and proportions. Most soils have four distinct horizons: O, A, B, and C.

Visual Connection Questions

[link] Soil compaction can result when soil is compressed by heavy machinery or even foot traffic. How might this compaction change the soil composition?

[link] The air content of the soil decreases.

[link] Which horizon is considered the topsoil, and which is considered the subsoil?

[link] The A horizon is the topsoil, and the B horizon is subsoil.

Review Questions

Which factors affect soil quality?

1. chemical composition
2. history of the soil
3. presence of living organisms and topography
4. all of the above

D

Soil particles that are 0.1 to 2 mm in diameter are called ______.

1. sand
2. silt
3. clay
4. loam

A

A soil consists of layers called ______ that taken together are called a ______.

1. soil profiles : horizon
2. horizons : soil profile
3. horizons : humus
4. humus : soil profile

B

What is the term used to describe the solid rock that lies beneath the soil?

1. sand
2. bedrock
3. clay
4. loam

B

Describe the main differences between a mineral soil and an organic soil.

A mineral soil forms from the weathering of rocks; it is inorganic material. An organic soil is formed from sedimentation; it mostly consists

of humus.

Name and briefly explain the factors that affect soil formation.

Parent material, climate, topography, biological factors, and time affect soil formation. Parent material is the material in which soils form. Climate describes how temperature, moisture, and wind cause different patterns of weathering, influencing the characteristics of the soil. Topography affects the characteristics and fertility of a soil. Biological factors include the presence of living organisms that greatly affect soil formation. Processes such as freezing and thawing may produce cracks in rocks; plant roots can penetrate these crevices and produce more fragmentation. Time affects soil because soil develops over long periods.

Describe how topography influences the characteristics and fertility of a soil.

Topography affects water runoff, which strips away parent material and affects plant growth. Steeps soils are more prone to erosion and may be thinner than soils that are on level surfaces.

Glossary

A horizon
: consists of a mixture of organic material with inorganic products of weathering

B horizon
: soil layer that is an accumulation of mostly fine material that has moved downward

bedrock
: solid rock that lies beneath the soil

C horizon
: layer of soil that contains the parent material, and the organic and inorganic material that is broken down to form soil; also known as the soil base

clay
: soil particles that are less than 0.002 mm in diameter

horizon
: soil layer with distinct physical and chemical properties, which differs from other layers depending on how and when it was formed

humus
: organic material of soil; made up of microorganisms, dead animals, and plants in varying stages of decay

loam
: soil that has no dominant particle size

mineral soil
: type of soil that is formed from the weathering of rocks and inorganic material; composed primarily of sand, silt, and clay

O horizon
: layer of soil with humus at the surface and decomposed vegetation at the base

organic soil
: type of soil that is formed from sedimentation; composed primarily of organic material

parent material
: organic and inorganic material in which soils form

rhizosphere
: area of soil affected by root secretions and microorganisms

sand
: soil particles between 0.1–2 mm in diameter

silt
: soil particles between 0.002 and 0.1 mm in diameter

soil profile

vertical section of a soil

soil
: outer loose layer that covers the surface of Earth

Nutritional Adaptations of Plants

By the end of this section, you will be able to do the following:

- Understand the nutritional adaptations of plants
- Describe mycorrhizae
- Explain nitrogen fixation

Plants obtain food in two different ways. Autotrophic plants can make their own food from inorganic raw materials, such as carbon dioxide and water, through photosynthesis in the presence of sunlight. Green plants are included in this group. Some plants, however, are heterotrophic: they are totally parasitic and lacking in chlorophyll. These plants, referred to as holo-parasitic plants, are unable to synthesize organic carbon and draw all of their nutrients from the host plant.

Plants may also enlist the help of microbial partners in nutrient acquisition. Particular species of bacteria and fungi have evolved along with certain plants to create a mutualistic symbiotic relationship with roots. This improves the nutrition of both the plant and the microbe. The formation of nodules in legume plants and mycorrhization can be considered among the nutritional adaptations of plants. However, these are not the only type of adaptations that we may find; many plants have other adaptations that allow them to thrive under specific

conditions.

Link to Learning

This video reviews basic concepts about photosynthesis. In the left panel, click each tab to select a topic for review.

Soybean roots contain (a) nitrogen-fixing nodules. Cells within the nodules are infected with *Bradyrhyzobium japonicum,* a rhizobia or “root-loving” bacterium. The bacteria are encased in (b) vesicles inside the cell, as can be seen in this transmission electron micrograph. (credit a: modification of work by USDA; credit b: modification of work by Louisa Howard, Dartmouth Electron Microscope Facility; scale-bar data from Matt Russell)

Nitrogen Fixation: Root and Bacteria Interactions

Nitrogen is an important macronutrient because it is part of nucleic acids and proteins. Atmospheric nitrogen, which is the diatomic molecule N_2, or dinitrogen, is the largest pool of nitrogen in terrestrial ecosystems. However, plants cannot take

advantage of this nitrogen because they do not have the necessary enzymes to convert it into biologically useful forms. However, nitrogen can be “fixed,” which means that it can be converted to ammonia (NH_3) through biological, physical, or chemical processes. As you have learned, biological nitrogen fixation (BNF) is the conversion of atmospheric nitrogen (N_2) into ammonia (NH_3), exclusively carried out by prokaryotes such as soil bacteria or cyanobacteria. Biological processes contribute 65 percent of the nitrogen used in agriculture. The following equation represents the process:

$$N_2 + 16\ ATP + 8\ e^- + 8\ H^+ \rightarrow 2NH_3 + 16\ ADP + 16\ Pi + H_2$$

The most important source of BNF is the symbiotic interaction between soil bacteria and legume plants, including many crops important to humans ([link]). The NH_3 resulting from fixation can be transported into plant tissue and incorporated into amino acids, which are then made into plant proteins. Some legume seeds, such as soybeans and peanuts, contain high levels of protein, and serve among the most important agricultural sources of protein in the world.

Visual Connection

Some common edible legumes—like (a) peanuts, (b) beans, and (c) chickpeas—are able to interact

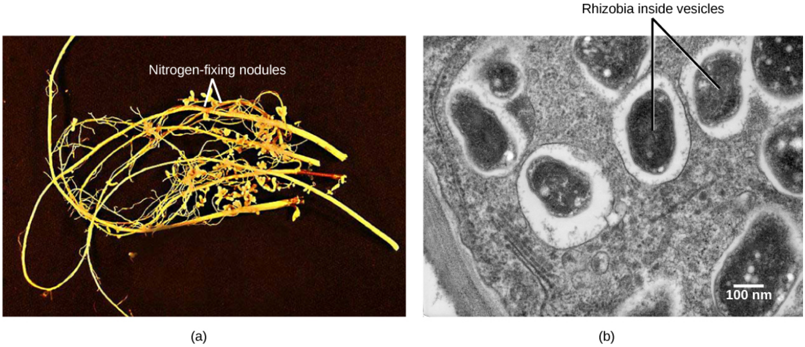
symbiotically with soil bacteria that fix nitrogen. (credit a: modification of work by Jules Clancy; credit b: modification of work by USDA)

(a) (b) (c)

Farmers often rotate corn (a cereal crop) and soy beans (a legume), planting a field with each crop in alternate seasons. What advantage might this crop rotation confer?

Soil bacteria, collectively called **rhizobia**, symbiotically interact with legume roots to form specialized structures called **nodules**, in which nitrogen fixation takes place. This process entails the reduction of atmospheric nitrogen to ammonia, by means of the enzyme **nitrogenase**. Therefore, using rhizobia is a natural and environmentally friendly way to fertilize plants, as opposed to chemical fertilization that uses a nonrenewable resource, such as natural gas. Through symbiotic nitrogen fixation, the plant benefits from using an endless source of nitrogen from the atmosphere. The process simultaneously contributes to soil fertility because the plant root system leaves behind some of the biologically available nitrogen. As in any symbiosis, both organisms benefit from the

interaction: the plant obtains ammonia, and bacteria obtain carbon compounds generated through photosynthesis, as well as a protected niche in which to grow ([link]).



Root tips proliferate in the presence of mycorrhizal infection, which appears as off-white fuzz in this image. (credit: modification of work by Nilsson et al., BMC Bioinformatics 2005)

Mycorrhizae: The Symbiotic Relationship between Fungi and Roots

A nutrient depletion zone can develop when there is rapid soil solution uptake, low nutrient concentration, low diffusion rate, or low soil moisture. These conditions are very common; therefore, most plants rely on fungi to facilitate the uptake of minerals from the soil. Fungi form symbiotic associations called mycorrhizae with plant roots, in which the fungi actually are integrated into the physical structure of the root. The fungi colonize the living root tissue during active plant growth.

Through mycorrhization, the plant obtains mainly phosphate and other minerals, such as zinc and copper, from the soil. The fungus obtains nutrients, such as sugars, from the plant root ([link]). Mycorrhizae help increase the surface area of the plant root system because hyphae, which are narrow, can spread beyond the nutrient depletion zone. Hyphae can grow into small soil pores that allow access to phosphorus that would otherwise be unavailable to the plant. The beneficial effect on the plant is best observed in poor soils. The benefit to fungi is that they can obtain up to 20 percent of the total carbon accessed by plants. Mycorrhizae functions as a physical barrier to pathogens. It also provides an induction of generalized host defense mechanisms, and sometimes involves production of antibiotic compounds by the fungi.

There are two types of mycorrhizae: ectomycorrhizae and endomycorrhizae. Ectomycorrhizae form an extensive dense sheath around the roots, called a mantle. Hyphae from the fungi extend from the mantle into the soil, which increases the surface area for water and mineral absorption. This type of mycorrhizae is found in forest trees, especially conifers, birches, and oaks. Endomycorrhizae, also called arbuscular mycorrhizae, do not form a dense sheath over the root. Instead, the fungal mycelium is embedded within the root tissue. Endomycorrhizae are found in the roots of more than 80 percent of terrestrial plants.

The dodder is a holoparasite that penetrates the host's vascular tissue and diverts nutrients for its own growth. Note that the vines of the dodder, which has white flowers, are beige. The dodder has no chlorophyll and cannot produce its own food. (credit: "Lalithamba"/Flickr) Saprophytes, like this Dutchmen's pipe (*Monotropa hypopitys)*, obtain their food from dead matter and do not have chlorophyll. (credit: modification of work by Iwona Erskine-Kellie)

Nutrients from Other Sources

Some plants cannot produce their own food and must obtain their nutrition from outside sources. This may occur with plants that are parasitic or saprophytic. Some plants are mutualistic symbionts,

epiphytes, or insectivorous.

Plant Parasites

A **parasitic plant** depends on its host for survival. Some parasitic plants have no leaves. An example of this is the dodder ([link]), which has a weak, cylindrical stem that coils around the host and forms suckers. From these suckers, cells invade the host stem and grow to connect with the vascular bundles of the host. The parasitic plant obtains water and nutrients through these connections. The plant is a total parasite (a holoparasite) because it is completely dependent on its host. Other parasitic plants (hemiparasites) are fully photosynthetic and only use the host for water and minerals. There are about 4,100 species of parasitic plants.

Saprophytes

A **saprophyte** is a plant that does not have chlorophyll and gets its food from dead matter, similar to bacteria and fungi (note that fungi are often called saprophytes, which is incorrect, because fungi are not plants). Plants like these use enzymes to convert organic food materials into simpler forms from which they can absorb nutrients ([link]). Most saprophytes do not directly digest dead matter: instead, they parasitize fungi that digest dead matter, or are mycorrhizal, ultimately obtaining photosynthate from a fungus that derived photosynthate from its host. Saprophytic plants are uncommon; only a few species are described.

Lichens, which often have symbiotic relationships

with other plants, can sometimes be found growing on trees. (credit: "benketaro"/Flickr) These epiphyte plants grow in the main greenhouse of the *Jardin des Plantes* in Paris.

Symbionts

A **symbiont** is a plant in a symbiotic relationship, with special adaptations such as mycorrhizae or nodule formation. Fungi also form symbiotic associations with cyanobacteria and green algae (called lichens). Lichens can sometimes be seen as colorful growths on the surface of rocks and trees ([link]). The algal partner (phycobiont) makes food autotrophically, some of which it shares with the fungus; the fungal partner (mycobiont) absorbs water and minerals from the environment, which are made available to the green alga. If one partner was separated from the other, they would both die.

Epiphytes

An **epiphyte** is a plant that grows on other plants, but is not dependent upon the other plant for nutrition ([link]). Epiphytes have two types of roots: clinging aerial roots, which absorb nutrients from humus that accumulates in the crevices of trees; and aerial roots, which absorb moisture from the atmosphere.

A Venus flytrap has specialized leaves to trap insects. (credit: "Selena N. B. H."/Flickr)

Insectivorous Plants

An **insectivorous** plant has specialized leaves to attract and digest insects. The Venus flytrap is popularly known for its insectivorous mode of nutrition, and has leaves that work as traps ([link]). The minerals it obtains from prey compensate for

those lacking in the boggy (low pH) soil of its native North Carolina coastal plains. There are three sensitive hairs in the center of each half of each leaf. The edges of each leaf are covered with long spines. Nectar secreted by the plant attracts flies to the leaf. When a fly touches the sensory hairs, the leaf immediately closes. Next, fluids and enzymes break down the prey and minerals are absorbed by the leaf. Since this plant is popular in the horticultural trade, it is threatened in its original habitat.

Section Summary

Atmospheric nitrogen is the largest pool of available nitrogen in terrestrial ecosystems. However, plants cannot use this nitrogen because they do not have

the necessary enzymes. Biological nitrogen fixation (BNF) is the conversion of atmospheric nitrogen to ammonia. The most important source of BNF is the symbiotic interaction between soil bacteria and legumes. The bacteria form nodules on the legume's roots in which nitrogen fixation takes place. Fungi form symbiotic associations (mycorrhizae) with plants, becoming integrated into the physical structure of the root. Through mycorrhization, the plant obtains minerals from the soil and the fungus obtains photosynthate from the plant root. Ectomycorrhizae form an extensive dense sheath around the root, while endomycorrhizae are embedded within the root tissue. Some plants—parasites, saprophytes, symbionts, epiphytes, and insectivores—have evolved adaptations to obtain their organic or mineral nutrition from various sources.

Visual Connection Questions

[link] Farmers often rotate corn (a cereal crop) and soy beans (a legume) planting a field with each crop in alternate seasons. What advantage might this crop rotation confer?

[link] Soybeans are able to fix nitrogen in their

roots, which are not harvested at the end of the growing season. The belowground nitrogen can be used in the next season by the corn.

Review Questions

Which process produces an inorganic compound that plants can easily use?

1. photosynthesis
2. nitrogen fixation
3. mycorrhization
4. Calvin cycle

B

Through mycorrhization, a plant obtains important nutrients such as _____.

1. phosphorus, zinc, and copper
2. phosphorus, zinc, and calcium
3. nickel, calcium, and zinc
4. all of the above

A

What term describes a plant that requires nutrition from a living host plant?

1. parasite
2. saprophyte
3. epiphyte
4. insectivorous

A

What is the term for the symbiotic association between fungi and cyanobacteria?

1. lichen
2. mycorrhizae
3. epiphyte
4. nitrogen-fixing nodule

A

Critical Thinking Questions

Why is biological nitrogen fixation an environmentally friendly way of fertilizing plants?

Because it is natural and does not require use of a nonrenewable resource, such as natural gas.

What is the main difference, from an energy point of view, between photosynthesis and biological nitrogen fixation?

Photosynthesis harvests and stores energy, whereas biological nitrogen fixation requires energy.

Why is a root nodule a nutritional adaptation of a plant?

A nodule results from the symbiosis between a plant and bacterium. Within nodules, the process of nitrogen fixation allows the plant to obtain nitrogen from the air.

Glossary

epiphyte
: plant that grows on other plants but is not dependent upon other plants for nutrition

insectivorous plant
: plant that has specialized leaves to attract and

digest insects

nitrogenase
: enzyme that is responsible for the reduction of atmospheric nitrogen to ammonia

nodules
: specialized structures that contain *Rhizobia* bacteria where nitrogen fixation takes place

parasitic plant
: plant that is dependent on its host for survival

rhizobia
: soil bacteria that symbiotically interact with legume roots to form nodules and fix nitrogen

saprophyte
: plant that does not have chlorophyll and gets its food from dead matter

symbiont
: plant in a symbiotic relationship with bacteria or fungi

Introduction
class="introduction" An arctic fox is a complex animal, well adapted to its environment. It changes coat color with the seasons, and has longer fur in winter to trap heat. (credit: modification of work by Keith Morehouse, USFWS)

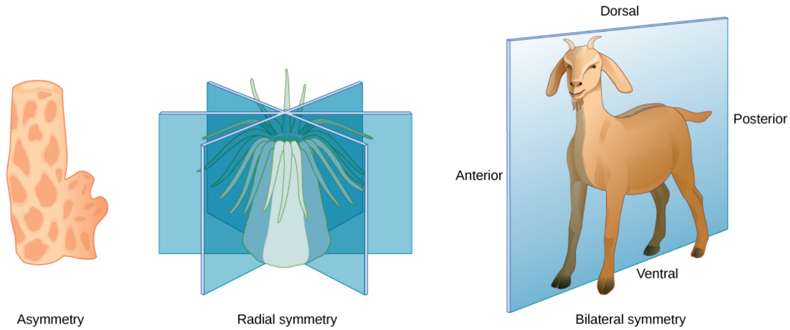
The arctic fox is an example of a complex animal that has adapted to its environment and illustrates the relationships between an animal's form and function. The structures of animals consist of primary tissues that make up more complex organs and organ systems. Homeostasis allows an animal to maintain a balance between its internal and external environments.

Animal Form and Function
By the end of this section, you will be able to do the following:

- Describe the various types of body plans that occur in animals
- Describe limits on animal size and shape
- Relate bioenergetics to body size, levels of activity, and the environment

Animals vary in form and function. From a sponge to a worm to a goat, an organism has a distinct body plan that limits its size and shape. Animals' bodies are also designed to interact with their environments, whether in the deep sea, a rainforest canopy, or the desert. Therefore, a large amount of information about the structure of an organism's body (anatomy) and the function of its cells, tissues and organs (physiology) can be learned by studying that organism's environment.
Animals exhibit different types of body symmetry. The sponge is asymmetrical, the sea anemone has radial symmetry, and the goat has bilateral symmetry.

Body Plans



Animal body plans follow set patterns related to symmetry. They are asymmetrical, radial, or bilateral in form as illustrated in [link]. **Asymmetrical** animals are animals with no pattern or symmetry; an example of an asymmetrical animal is a sponge. Radial symmetry, as illustrated in [link], describes when an animal has an up-and-down orientation: any plane cut along its longitudinal axis through the organism produces equal halves, but not a definite right or left side. This plan is found mostly in aquatic animals, especially organisms that attach themselves to a base, like a rock or a boat, and extract their food from the surrounding water as it flows around the organism. Bilateral symmetry is illustrated in the same figure by a goat. The goat also has an upper and lower component to it, but a plane cut from front to back separates the animal into definite right and left sides. Additional terms used when describing positions in the body are anterior (front), posterior (rear), dorsal (toward the back), and ventral (toward the stomach). Bilateral symmetry is found in both land-based and aquatic animals; it

enables a high level of mobility.

Apodemes are ingrowths on arthropod exoskeletons to which muscles attach. The apodemes on this crab leg are located above and below the fulcrum of the claw. Contraction of muscles attached to the apodemes pulls the claw closed.

Limits on Animal Size and Shape

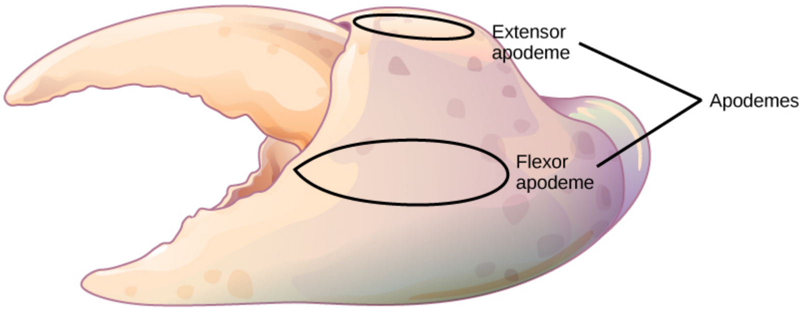
Animals with bilateral symmetry that live in water tend to have a **fusiform** shape: this is a tubular shaped body that is tapered at both ends. This shape decreases the drag on the body as it moves through water and allows the animal to swim at high speeds. [link] lists the maximum speed of various animals. Certain types of sharks can swim at fifty kilometers per hour and some dolphins at 32 to 40 kilometers per hour. Land animals frequently travel faster, although the tortoise and snail are significantly slower than cheetahs. Another difference in the adaptations of aquatic and land-dwelling organisms is that aquatic organisms are constrained in shape by the forces of drag in the water since water has higher viscosity than air. On the other hand, land-dwelling organisms are constrained mainly by gravity, and drag is relatively unimportant. For example, most adaptations in birds are for gravity not for drag.

Maximum Speed of Assorted Land & Marine Animals		
Animal	**Speed (kmh)**	**Speed (mph)**
Cheetah	113	70
Quarter horse	77	48
Fox	68	42
Shortfin mako shark	50	31
Domestic house cat	48	30
Human	45	28
Dolphin	32–40	20–25
Mouse	13	8
Snail	0.05	0.03

Most animals have an exoskeleton, including insects, spiders, scorpions, horseshoe crabs, centipedes, and crustaceans. Scientists estimate that, of insects alone, there are over 30 million species on our planet. The exoskeleton is a hard covering or shell that provides benefits to the animal, such as protection against damage from predators and from water loss (for land animals); it also provides for the attachments of muscles.

As the tough and resistant outer cover of an arthropod, the exoskeleton may be constructed of a tough polymer such as chitin and is often biomineralized with materials such as calcium

carbonate. This is fused to the animal's epidermis. Ingrowths of the exoskeleton, called **apodemes**, function as attachment sites for muscles, similar to tendons in more advanced animals ([link]). In order to grow, the animal must first synthesize a new exoskeleton underneath the old one and then shed or molt the original covering. This limits the animal's ability to grow continually, and may limit the individual's ability to mature if molting does not occur at the proper time. The thickness of the exoskeleton must be increased significantly to accommodate any increase in weight. It is estimated that a doubling of body size increases body weight by a factor of eight. The increasing thickness of the chitin necessary to support this weight limits most animals with an exoskeleton to a relatively small size. The same principles apply to endoskeletons, but they are more efficient because muscles are attached on the outside, making it easier to compensate for increased mass.



An animal with an endoskeleton has its size determined by the amount of skeletal system it needs in order to support the other tissues and the

amount of muscle it needs for movement. As the body size increases, both bone and muscle mass increase. The speed achievable by the animal is a balance between its overall size and the bone and muscle that provide support and movement.

Limiting Effects of Diffusion on Size and Development

The exchange of nutrients and wastes between a cell and its watery environment occurs through the process of diffusion. All living cells are bathed in liquid, whether they are in a single-celled organism or a multicellular one. Diffusion is effective over a specific distance and limits the size that an individual cell can attain. If a cell is a single-celled microorganism, such as an amoeba, it can satisfy all of its nutrient and waste needs through diffusion. If the cell is too large, then diffusion is ineffective and the center of the cell does not receive adequate nutrients nor is it able to effectively dispel its waste.

An important concept in understanding how efficient diffusion is as a means of transport is the surface to volume ratio. Recall that any three-dimensional object has a surface area and volume; the ratio of these two quantities is the surface-to-volume ratio. Consider a cell shaped like a perfect sphere: it has a surface area of $4\pi r^2$, and a volume of $(4/3)\pi r^3$. The surface-to-volume ratio of a sphere

is 3/r; as the cell gets bigger, its surface to volume ratio decreases, making diffusion less efficient. The larger the size of the sphere, or animal, the less surface area for diffusion it possesses.

The solution to producing larger organisms is for them to become multicellular. Specialization occurs in complex organisms, allowing cells to become more efficient at doing fewer tasks. For example, circulatory systems bring nutrients and remove waste, while respiratory systems provide oxygen for the cells and remove carbon dioxide from them. Other organ systems have developed further specialization of cells and tissues and efficiently control body functions. Moreover, surface-to-volume ratio applies to other areas of animal development, such as the relationship between muscle mass and cross-sectional surface area in supporting skeletons, and in the relationship between muscle mass and the generation of dissipation of heat.

Link to Learning

Visit this interactive site to see an entire animal (a zebrafish embryo) at the cellular and sub-cellular level. Use the zoom and navigation functions for a virtual nanoscopy exploration.

The mouse has a much higher metabolic rate than the elephant. (credit “mouse”: modification of work by Magnus Kjaergaard; credit “elephant”: modification of work by “TheLizardQueen”/Flickr)

Animal Bioenergetics

All animals must obtain their energy from food they ingest or absorb. These nutrients are converted to adenosine triphosphate (ATP) for short-term storage and use by all cells. Some animals store energy for slightly longer times as glycogen, and others store energy for much longer times in the form of triglycerides housed in specialized adipose tissues. No energy system is one hundred percent efficient, and an animal’s metabolism produces waste energy in the form of heat. If an animal can conserve that heat and maintain a relatively constant body temperature, it is classified as a warm-blooded animal and called an **endotherm**. The insulation used to conserve the body heat comes in the forms of fur, fat, or feathers. The absence of insulation in **ectothermic** animals increases their dependence on the environment for body heat.

The amount of energy expended by an animal over a specific time is called its metabolic rate. The rate is measured variously in joules, calories, or kilocalories (1000 calories). Carbohydrates and proteins contain about 4.5 to 5 kcal/g, and fat contains about 9 kcal/g. Metabolic rate is estimated

as the **basal metabolic rate (BMR)** in endothermic animals at rest and as the **standard metabolic rate (SMR)** in ectotherms. Human males have a BMR of 1600 to 1800 kcal/day, and human females have a BMR of 1300 to 1500 kcal/day. Even with insulation, endothermal animals require extensive amounts of energy to maintain a constant body temperature. An ectotherm such as an alligator has an SMR of 60 kcal/day.

Energy Requirements Related to Body Size

Smaller endothermic animals have a greater surface area for their mass than larger ones ([link]). Therefore, smaller animals lose heat at a faster rate than larger animals and require more energy to maintain a constant internal temperature. This results in a smaller endothermic animal having a higher BMR, per body weight, than a larger endothermic animal.

Species		
Mass	35 g	4,500,000 g
Metabolic rate	890 mm^3 0_2/g body mass/hr	75 mm^3 0_2/g body mass/hr

Energy Requirements Related to Levels of Activity

The more active an animal is, the more energy is needed to maintain that activity, and the higher its BMR or SMR. The average daily rate of energy consumption is about two to four times an animal's BMR or SMR. Humans are more sedentary than most animals and have an average daily rate of only 1.5 times the BMR. The diet of an endothermic animal is determined by its BMR. For example: the type of grasses, leaves, or shrubs that an herbivore eats affects the number of calories that it takes in. The relative caloric content of herbivore foods, in descending order, is tall grasses > legumes > short grasses > forbs (any broad-leaved plant, not a grass) > subshrubs > annuals/biennials.

Energy Requirements Related to Environment

Animals adapt to extremes of temperature or food availability through torpor. **Torpor** is a process that leads to a decrease in activity and metabolism and allows animals to survive adverse conditions. Torpor can be used by animals for long periods, such as entering a state of **hibernation** during the winter months, in which case it enables them to maintain a reduced body temperature. During hibernation, ground squirrels can achieve an abdominal temperature of 0° C (32° F), while a bear's internal temperature is maintained higher at about 37° C (99° F).

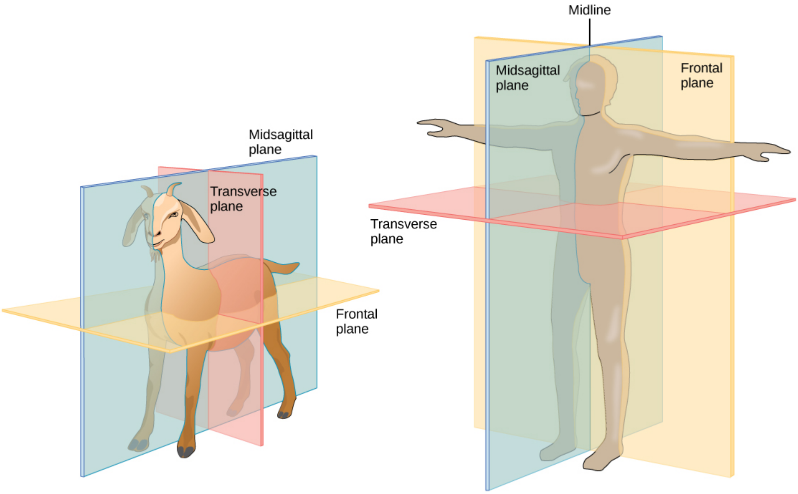
If torpor occurs during the summer months with

high temperatures and little water, it is called **estivation**. Some desert animals use this to survive the harshest months of the year. Torpor can occur on a daily basis; this is seen in bats and hummingbirds. While endothermy is limited in smaller animals by surface to volume ratio, some organisms can be smaller and still be endotherms because they employ daily torpor during the part of the day that is coldest. This allows them to conserve energy during the colder parts of the day, when they consume more energy to maintain their body temperature.

Shown are the planes of a quadrupedal goat and a bipedal human. The midsagittal plane divides the body exactly in half, into right and left portions. The frontal plane divides the front and back, and the transverse plane divides the body into upper and lower portions. Vertebrate animals have two major body cavities. The dorsal cavity, indicated in green, contains the cranial and the spinal cavity. The ventral cavity, indicated in yellow, contains the thoracic cavity and the abdominopelvic cavity. The thoracic cavity is separated from the abdominopelvic cavity by the diaphragm. The thoracic cavity is separated into the abdominal cavity and the pelvic cavity by an imaginary line parallel to the pelvis bones. (credit: modification of work by NCI)

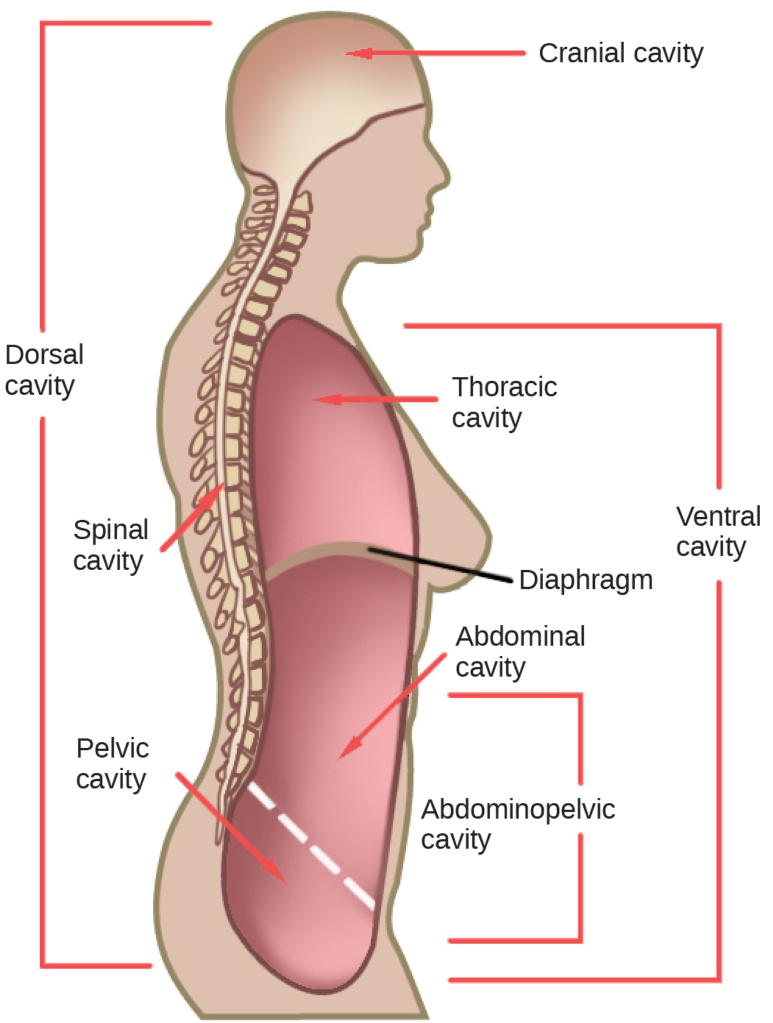
Animal Body Planes and Cavities

A standing vertebrate animal can be divided by several planes. A **sagittal plane** divides the body into right and left portions. A **midsagittal plane** divides the body exactly in the middle, making two equal right and left halves. A **frontal plane** (also called a coronal plane) separates the front from the back. A **transverse plane** (or, horizontal plane) divides the animal into upper and lower portions. This is sometimes called a cross section, and, if the transverse cut is at an angle, it is called an oblique plane. [link] illustrates these planes on a goat (a four-legged animal) and a human being.



Vertebrate animals have a number of defined body cavities, as illustrated in [link]. Two of these are major cavities that contain smaller cavities within them. The **dorsal cavity** contains the cranial and the vertebral (or spinal) cavities. The **ventral cavity** contains the thoracic cavity, which in turn contains the pleural cavity around the lungs and the

pericardial cavity, which surrounds the heart. The ventral cavity also contains the abdominopelvic cavity, which can be separated into the abdominal and the pelvic cavities.



Career Connections

Physical Anthropologist

Physical anthropologists study the adaption, variability, and evolution of human beings, plus their living and fossil relatives. They can work in a variety of settings, although most will have an academic appointment at a university, usually in an anthropology department or a biology, genetics, or zoology department.

Nonacademic positions are available in the automotive and aerospace industries where the focus is on human size, shape, and anatomy. Research by these professionals might range from studies of how the human body reacts to car crashes to exploring how to make seats more comfortable. Other nonacademic positions can be obtained in museums of natural history, anthropology, archaeology, or science and technology. These positions involve educating students from grade school through graduate school. Physical anthropologists serve as education coordinators, collection managers, writers for museum publications, and as administrators. Zoos employ these professionals, especially if they have an expertise in primate biology; they work in collection management and captive breeding programs for endangered species. Forensic science utilizes physical anthropology expertise in identifying human and animal remains, assisting in determining the cause of death, and for expert testimony in trials.

Section Summary

Animal bodies come in a variety of sizes and shapes. Limits on animal size and shape include impacts to their movement. Diffusion affects their size and development. Bioenergetics describes how animals use and obtain energy in relation to their body size, activity level, and environment.

Review Questions

Which type of animal maintains a constant internal body temperature?

1. endotherm
2. ectotherm
3. coelomate
4. mesoderm

A

The symmetry found in animals that move swiftly is ______.

1. radial
2. bilateral
3. sequential
4. interrupted

B

What term describes the condition of a desert mouse that lowers its metabolic rate and "sleeps" during the hot day?

1. turgid
2. hibernation
3. estivation
4. normal sleep pattern

C

A plane that divides an animal into equal right and left portions is ______.

1. diagonal
2. midsagittal
3. coronal
4. transverse

B

A plane that divides an animal into dorsal and ventral portions is _____.

1. sagittal
2. midsagittal
3. coronal
4. transverse

D

The pleural cavity is a part of which cavity?

1. dorsal cavity
2. thoracic cavity
3. abdominal cavity
4. pericardial cavity

B

How could the increasing global temperature associated with climate change impact ectotherms?

1. Ectotherm diversity will decrease in cool regions.
2. Ectotherms will be able to be active all day in the tropics.
3. Ectotherms will have to expend more

energy to cool their body temperatures.
4. Ectotherms will be able to expand into new habitats.

D

Although most animals are bilaterally symmetrical, a few exhibit radial symmetry. What is an advantage of radial symmetry?

1. It confuses predators.
2. It allows the animal to gather food from all sides.
3. It allows the animal to undergo rapid, purposeful movement in any direction.
4. It lets an animal use its dorsal surface to sense its environment.

B

Critical Thinking Questions

How does diffusion limit the size of an organism? How is this counteracted?

Diffusion is effective over a very short distance. If a cell exceeds this distance in its size, the center of the cell cannot get adequate nutrients nor can it expel enough waste to survive. To compensate for this, cells can loosely adhere to each other in a liquid medium, or develop into multi-celled organisms that use circulatory and respiratory systems to deliver nutrients and remove wastes.

What is the relationship between BMR and body size? Why?

Basal Metabolic Rate is an expression of the metabolic processes that occur to maintain an individual's functioning and body temperature. Smaller bodied animals have a relatively large surface area compared to a much larger animal. The large animal's large surface area leads to increased heat loss that the animal must compensate for, resulting in a higher BMR. A small animal, having less relative surface area, does not lose as much heat and has a correspondingly lower BMR.

Explain how using an open circulatory system constrains the size of animals.

In an open circulatory system, the heart(s) pump blood into an open cavity, bathing the tissues. As the blood diffuses through the tissue space, it delivers nutrients in exchange for receiving metabolic wastes. The blood then diffuses back to the heart to be pumped again. However, since this system relies on diffusion, the size of animals that use an open circulatory system is limited to fairly small volumes so that the blood can diffuse rapidly enough to efficiently exchange molecules with the tissues.

Describe one key environmental constraint for ectotherms and one for endotherms. Why are they limited by different factors?

Endotherms are constrained by the availability of food sources in the environment, while the temperature range in a geographic area limits ectotherms. The difference in how the two groups maintain their body temperature determines the key constraint for each group.

Glossary

apodeme
: ingrowth of an animal's exoskeleton that functions as an attachment site for muscles

asymmetrical
: describes animals with no axis of symmetry in their body pattern

basal metabolic rate (BMR)
: metabolic rate at rest in endothermic animals

dorsal cavity
: body cavity on the posterior or back portion of an animal; includes the cranial and vertebral cavities

ectotherm
: animal incapable of maintaining a relatively constant internal body temperature

endotherm
: animal capable of maintaining a relatively constant internal body temperature

estivation
: torpor in response to extremely high temperatures and low water availability

frontal (coronal) plane
: plane cutting through an animal separating the individual into front and back portions

fusiform
: animal body shape that is tubular and tapered at both ends

hibernation

torpor over a long period of time, such as a winter

midsagittal plane
: plane cutting through an animal separating the individual into even right and left sides

sagittal plane
: plane cutting through an animal separating the individual into right and left sides

standard metabolic rate (SMR)
: metabolic rate at rest in ectothermic animals

torpor
: decrease in activity and metabolism that allows an animal to survive adverse conditions

transverse (horizontal) plane
: plane cutting through an animal separating the individual into upper and lower portions

ventral cavity
: body cavity on the anterior or front portion of an animal that includes the thoracic cavities and the abdominopelvic cavities

Animal Primary Tissues

By the end of this section, you will be able to do the following:

- Describe epithelial tissues
- Discuss the different types of connective tissues in animals
- Describe three types of muscle tissues
- Describe nervous tissue

The tissues of multicellular, complex animals are four primary types: epithelial, connective, muscle, and nervous. Recall that tissues are groups of similar cells (cells carrying out related functions). These tissues combine to form organs—like the skin or kidney—that have specific, specialized functions within the body. Organs are organized into organ systems to perform functions; examples include the circulatory system, which consists of the heart and blood vessels, and the digestive system, consisting of several organs, including the stomach, intestines, liver, and pancreas. Organ systems come together to create an entire organism.

Squamous epithelia cells (a) have a slightly irregular shape, and a small, centrally located nucleus. These cells can be stratified into layers, as in (b) this human cervix specimen. (credit b: modification of work by Ed Uthman; scale-bar data from Matt Russell) Simple cuboidal epithelial cells line tubules in the mammalian kidney, where they are involved in filtering the blood. Simple columnar

epithelial cells absorb material from the digestive tract. Goblet cells secrete mucus into the digestive tract lumen.Pseudostratified columnar epithelia line the respiratory tract. They exist in one layer, but the arrangement of nuclei at different levels makes it appear that there is more than one layer. Goblet cells interspersed between the columnar epithelial cells secrete mucus into the respiratory tract.

Epithelial Tissues

Epithelial tissues cover the outside of organs and structures in the body and line the lumens of organs in a single layer or multiple layers of cells. The types of epithelia are classified by the shapes of cells present and the number of layers of cells. Epithelia composed of a single layer of cells is called **simple epithelia**; epithelial tissue composed of multiple layers is called **stratified epithelia**. [link] summarizes the different types of epithelial tissues.

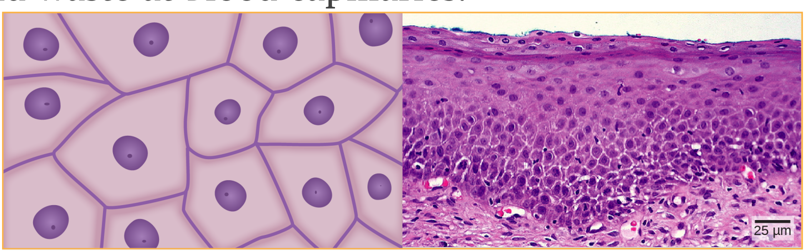
Different Types of Epithelial Tissues

Cell shape	Description	Location
squamous	flat, irregular	simple: lung

	round shape	alveoli, capillaries; stratified: skin, mouth, vagina
cuboidal	cube shaped, central nucleus	glands, renal tubules
columnar	tall, narrow, nucleus toward base; tall, narrow, nucleus along cell	simple: digestive tract; pseudostratified: respiratory tract
transitional	round, simple but appear stratified	urinary bladder

Squamous Epithelia

Squamous epithelial cells are generally round, flat, and have a small, centrally located nucleus. The cell outline is slightly irregular, and cells fit together to form a covering or lining. When the cells are arranged in a single layer (simple epithelia), they facilitate diffusion in tissues, such as the areas of gas exchange in the lungs and the exchange of nutrients and waste at blood capillaries.

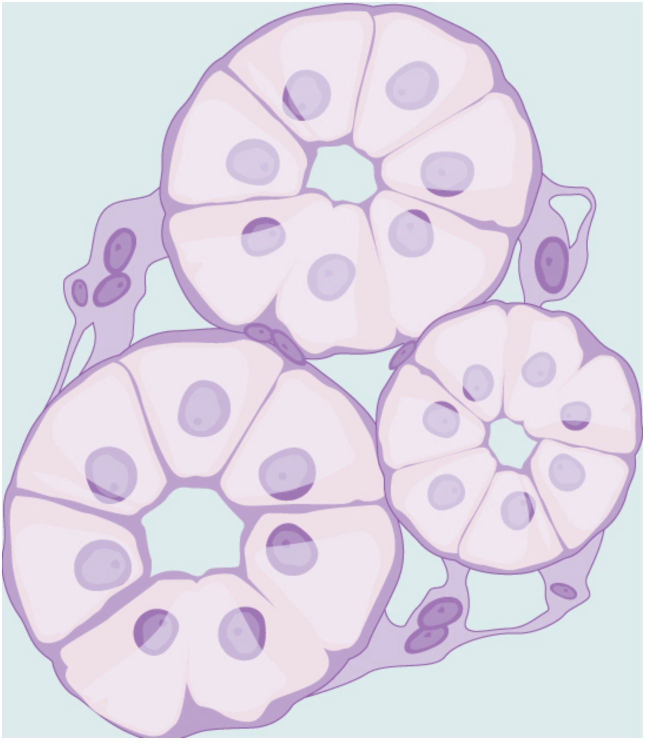


(a) (b)

[link]**a** illustrates a layer of squamous cells with their membranes joined together to form an epithelium. Image [link]**b** illustrates squamous epithelial cells arranged in stratified layers, where protection is needed on the body from outside abrasion and damage. This is called a stratified squamous epithelium and occurs in the skin and in tissues lining the mouth and vagina.

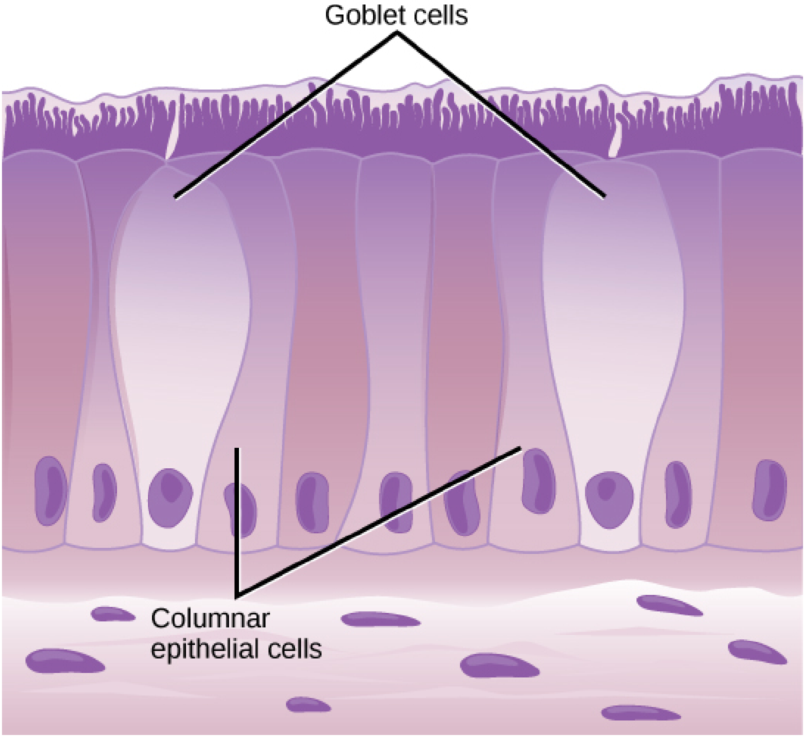
Cuboidal Epithelia

Cuboidal epithelial cells, shown in [link], are cube-shaped with a single, central nucleus. They are most commonly found in a single layer representing a simple epithelia in glandular tissues throughout the body where they prepare and secrete glandular material. They are also found in the walls of tubules and in the ducts of the kidney and liver.



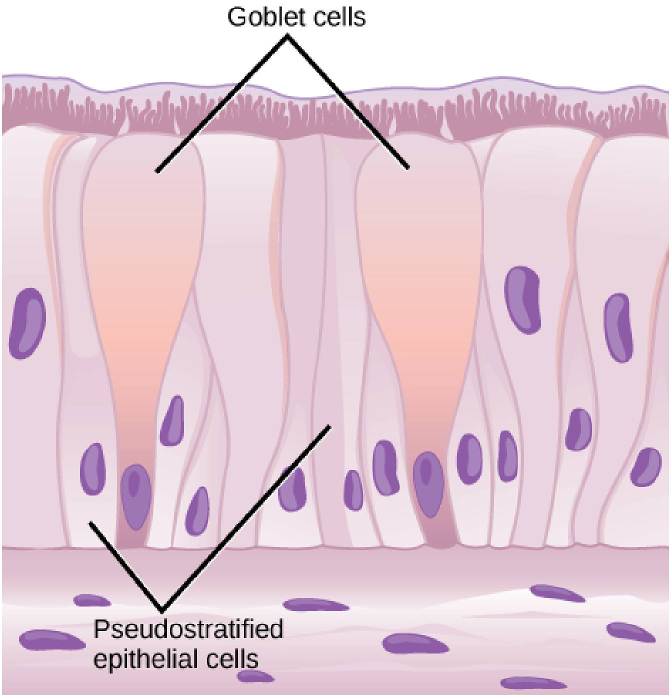
Columnar Epithelia

Columnar epithelial cells are taller than they are wide: they resemble a stack of columns in an epithelial layer, and are most commonly found in a single-layer arrangement. The nuclei of columnar epithelial cells in the digestive tract appear to be lined up at the base of the cells, as illustrated in [link]. These cells absorb material from the lumen of the digestive tract and prepare it for entry into the body through the circulatory and lymphatic systems.



Columnar epithelial cells lining the respiratory tract appear to be stratified. However, each cell is attached to the base membrane of the tissue and, therefore, they are simple tissues. The nuclei are arranged at different levels in the layer of cells, making it appear as though there is more than one layer, as seen in [link]. This is called **pseudostratified**, columnar epithelia. This cellular covering has cilia at the apical, or free, surface of the cells. The cilia enhance the movement of mucus and trapped particles out of the respiratory tract, helping to protect the system from invasive microorganisms and harmful material that has been

breathed into the body. Goblet cells are interspersed in some tissues (such as the lining of the trachea). The goblet cells contain mucus that traps irritants, which in the case of the trachea keep these irritants from getting into the lungs.



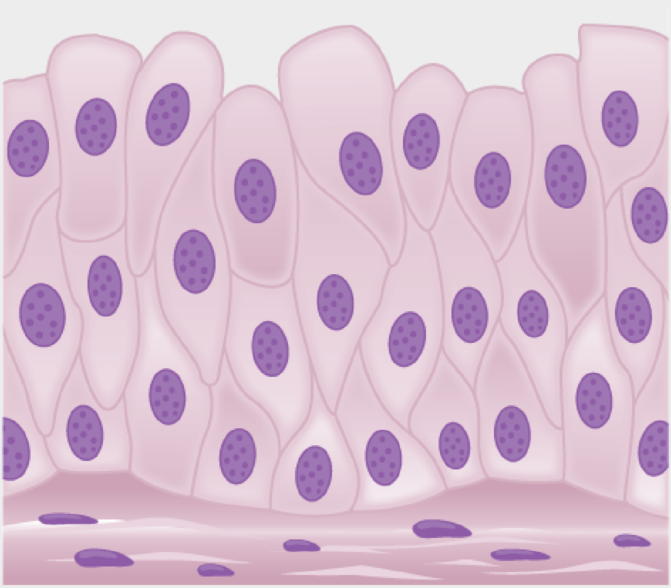
Transitional Epithelia

Transitional or uroepithelial cells appear only in the urinary system, primarily in the bladder and ureter. These cells are arranged in a stratified layer, but they have the capability of appearing to pile up on top of each other in a relaxed, empty bladder, as illustrated in [link]. As the urinary bladder fills, the epithelial layer unfolds and expands to hold the

volume of urine introduced into it. As the bladder fills, it expands and the lining becomes thinner. In other words, the tissue transitions from thick to thin.

Visual Connection

Transitional epithelia of the urinary bladder undergo changes in thickness depending on how full the bladder is.



Which of the following statements about types of epithelial cells is false?

1. Simple columnar epithelial cells line the tissue of the lung.
2. Simple cuboidal epithelial cells are involved in

the filtering of blood in the kidney.

3. Pseudostratisfied columnar epithilia occur in a single layer, but the arrangement of nuclei makes it appear that more than one layer is present.
4. Transitional epithelia change in thickness depending on how full the bladder is.

Loose connective tissue is composed of loosely woven collagen and elastic fibers. The fibers and other components of the connective tissue matrix are secreted by fibroblasts. Fibrous connective tissue from the tendon has strands of collagen fibers lined up in parallel. Hyaline cartilage consists of a matrix with cells called chondrocytes embedded in it. The chondrocytes exist in cavities in the matrix called lacunae. (a) Compact bone is a dense matrix on the outer surface of bone. Spongy bone, inside the compact bone, is porous with web-like trabeculae. (b) Compact bone is organized into rings called osteons. Blood vessels, nerves, and lymphatic vessels are found in the central Haversian canal. Rings of lamellae surround the Haversian canal. Between the lamellae are cavities called lacunae. Canaliculi are microchannels connecting the lacunae together. (c) Osteoblasts surround the exterior of the bone. Osteoclasts bore tunnels into the bone and osteocytes are found in the lacunae. Adipose is a connective tissue is made up of cells called

adipocytes. Adipocytes have small nuclei localized at the cell edge. Blood is a connective tissue that has a fluid matrix, called plasma, and no fibers. Erythrocytes (red blood cells), the predominant cell type, are involved in the transport of oxygen and carbon dioxide. Also present are various leukocytes (white blood cells) involved in immune response.

Connective Tissues

Connective tissues are made up of a matrix consisting of living cells and a nonliving substance, called the ground substance. The ground substance is made of an organic substance (usually a protein) and an inorganic substance (usually a mineral or water). The principal cell of connective tissues is the fibroblast. This cell makes the fibers found in nearly all of the connective tissues. Fibroblasts are motile, able to carry out mitosis, and can synthesize whichever connective tissue is needed. Macrophages, lymphocytes, and, occasionally, leukocytes can be found in some of the tissues. Some tissues have specialized cells that are not found in the others. The **matrix** in connective tissues gives the tissue its density. When a connective tissue has a high concentration of cells or fibers, it has proportionally a less dense matrix.

The organic portion or protein fibers found in connective tissues are either collagen, elastic, or reticular fibers. Collagen fibers provide strength to

the tissue, preventing it from being torn or separated from the surrounding tissues. Elastic fibers are made of the protein elastin; this fiber can stretch to one and one half of its length and return to its original size and shape. Elastic fibers provide flexibility to the tissues. Reticular fibers are the third type of protein fiber found in connective tissues. This fiber consists of thin strands of collagen that form a network of fibers to support the tissue and other organs to which it is connected. The various types of connective tissues, the types of cells and fibers they are made of, and sample locations of the tissues is summarized in [link].

Connective Tissues

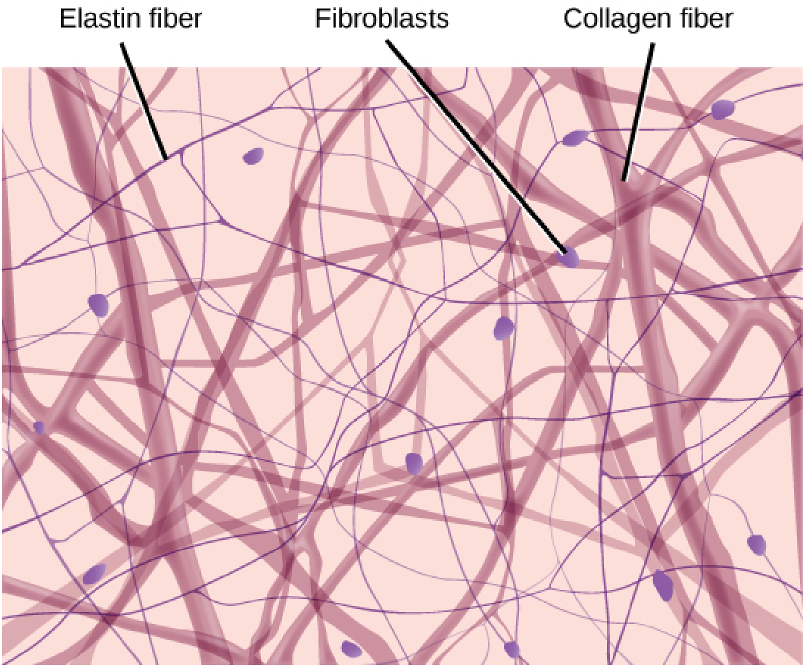
Tissue	Cells	Fibers	Location
loose/areolar	fibroblasts, macrophages, some lymphocytes, some neutrophils	few: collagen, elastic, reticular	around blood vessels; anchors epithelia
dense, fibrous connective tissue	fibroblasts, macrophages	mostly collagen	irregular: skin; regular: tendons, ligaments

cartilage	chondrocytes, chondroblasts	hyaline: few: collagen fibrocartilage: large amount of collagen	shark skeleton, fetal bones, human ears, intervertebral discs
bone	osteoblasts, osteocytes, osteoclasts	some: collagen, elastic	vertebrate skeletons
adipose	adipocytes	few	adipose (fat)
blood	red blood cells, white blood cells	none	blood

Loose/Areolar Connective Tissue

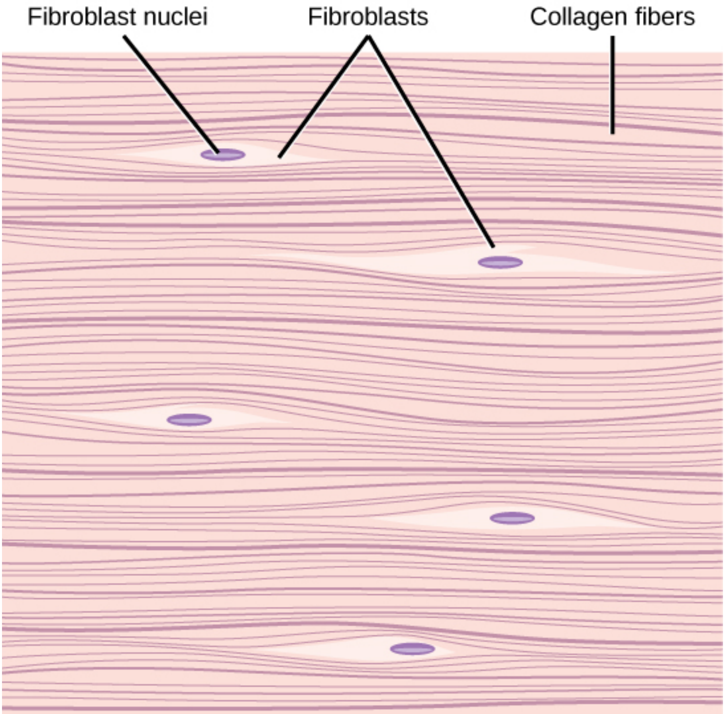
Loose connective tissue, also called areolar connective tissue, has a sampling of all of the components of a connective tissue. As illustrated in [link], loose connective tissue has some fibroblasts; macrophages are present as well. Collagen fibers are relatively wide and stain a light pink, while elastic fibers are thin and stain dark blue to black. The space between the formed elements of the tissue is filled with the matrix. The material in the connective tissue gives it a loose consistency similar to a cotton ball that has been pulled apart. Loose connective tissue is found around every blood vessel and helps to keep the vessel in place. The tissue is also found around and between most body organs. In summary, areolar tissue is tough, yet flexible, and

comprises membranes.



Fibrous Connective Tissue

Fibrous connective tissues contain large amounts of collagen fibers and few cells or matrix material. The fibers can be arranged irregularly or regularly with the strands lined up in parallel. Irregularly arranged fibrous connective tissues are found in areas of the body where stress occurs from all directions, such as the dermis of the skin. Regular fibrous connective tissue, shown in [link], is found in tendons (which connect muscles to bones) and ligaments (which connect bones to bones).

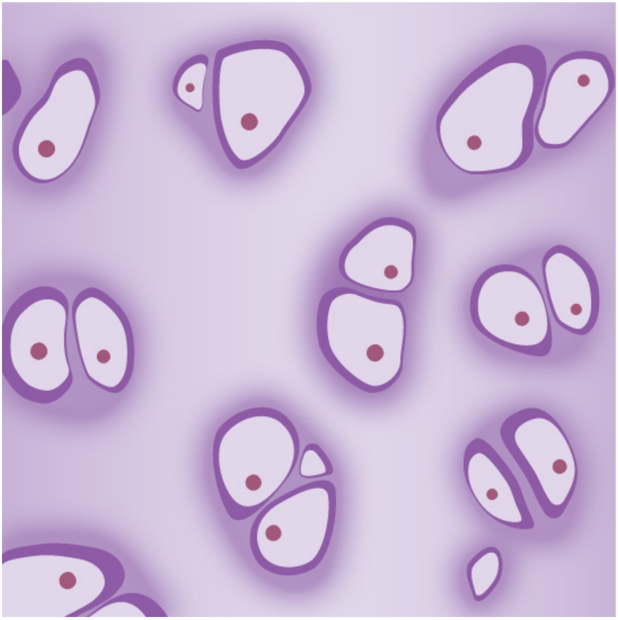


Cartilage

Cartilage is a connective tissue with a large amount of the matrix and variable amounts of fibers. The cells, called **chondrocytes**, make the matrix and fibers of the tissue. Chondrocytes are found in spaces within the tissue called **lacunae**.

A cartilage with few collagen and elastic fibers is hyaline cartilage, illustrated in [link]. The lacunae are randomly scattered throughout the tissue and the matrix takes on a milky or scrubbed appearance with routine histological stains. Sharks have cartilaginous skeletons, as does nearly the entire

human skeleton during a specific pre-birth developmental stage. A remnant of this cartilage persists in the outer portion of the human nose. Hyaline cartilage is also found at the ends of long bones, reducing friction and cushioning the articulations of these bones.



Elastic cartilage has a large amount of elastic fibers, giving it tremendous flexibility. The ears of most vertebrate animals contain this cartilage as do portions of the larynx, or voice box. Fibrocartilage contains a large amount of collagen fibers, giving the tissue tremendous strength. Fibrocartilage comprises the intervertebral discs in vertebrate animals. Hyaline cartilage found in movable joints such as the knee and shoulder becomes damaged as a result of age or trauma. Damaged hyaline cartilage is replaced by fibrocartilage and results in the joints

becoming “stiff.”

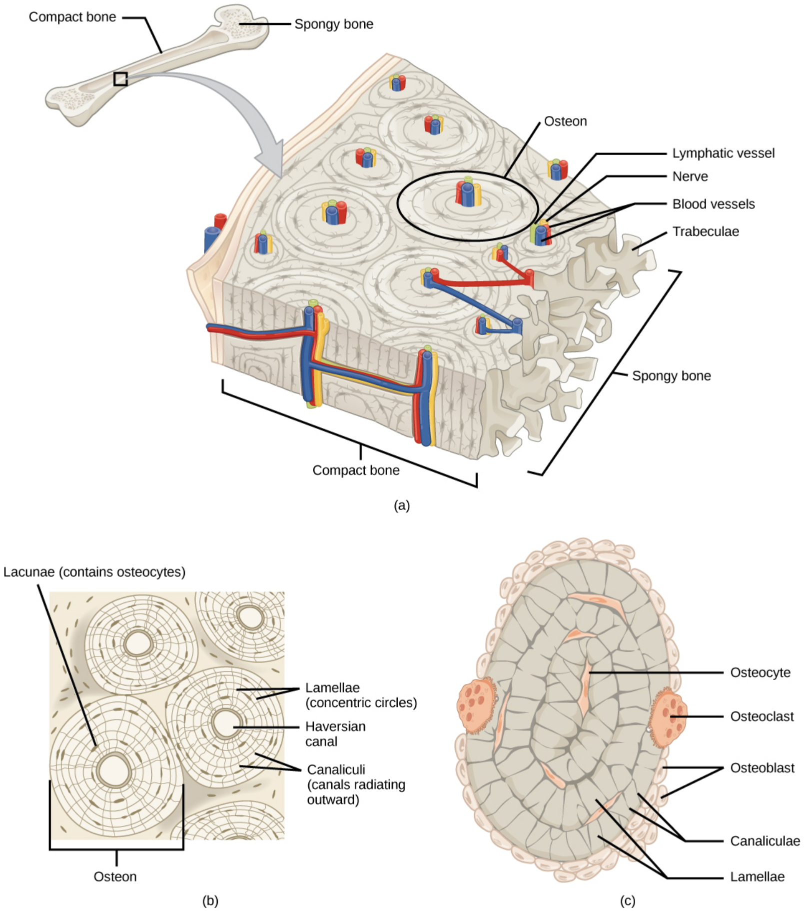
Bone

Bone, or osseous tissue, is a connective tissue that has a large amount of two different types of matrix material. The organic matrix is similar to the matrix material found in other connective tissues, including some amount of collagen and elastic fibers. This gives strength and flexibility to the tissue. The inorganic matrix consists of mineral salts—mostly calcium salts—that give the tissue hardness. Without adequate organic material in the matrix, the tissue breaks; without adequate inorganic material in the matrix, the tissue bends.

There are three types of cells in bone: osteoblasts, osteocytes, and osteoclasts. Osteoblasts are active in making bone for growth and remodeling. Osteoblasts deposit bone material into the matrix and, after the matrix surrounds them, they continue to live, but in a reduced metabolic state as osteocytes. Osteocytes are found in lacunae of the bone. Osteoclasts are active in breaking down bone for bone remodeling, and they provide access to calcium stored in tissues. Osteoclasts are usually found on the surface of the tissue.

Bone can be divided into two types: compact and spongy. Compact bone is found in the shaft (or diaphysis) of a long bone and the surface of the flat

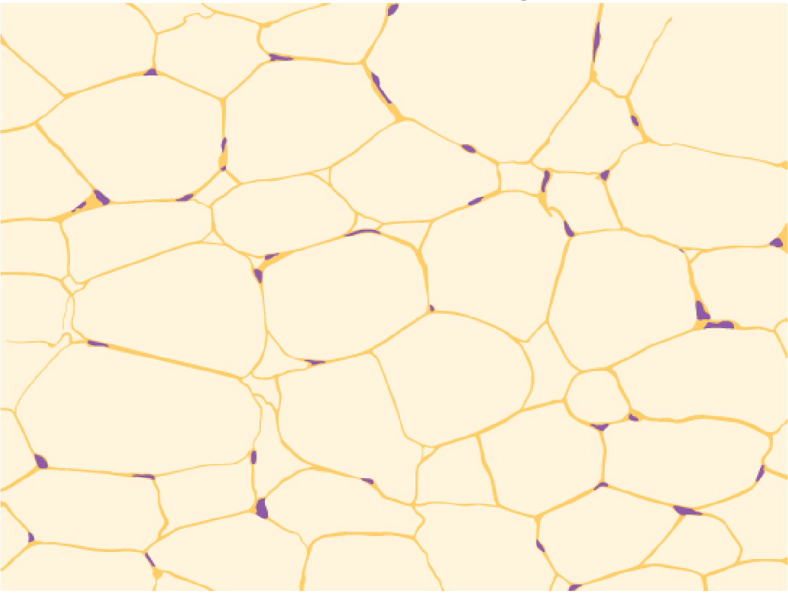
bones, while spongy bone is found in the end (or epiphysis) of a long bone. Compact bone is organized into subunits called **osteons**, as illustrated in [link]. A blood vessel and a nerve are found in the center of the structure within the Haversian canal, with radiating circles of lacunae around it known as lamellae. The wavy lines seen between the lacunae are microchannels called **canaliculi**; they connect the lacunae to aid diffusion between the cells. Spongy bone is made of tiny plates called **trabeculae**; these plates serve as struts to give the spongy bone strength. Over time, these plates can break causing the bone to become less resilient. Bone tissue forms the internal skeleton of vertebrate animals, providing structure to the animal and points of attachment for tendons.



Adipose Tissue

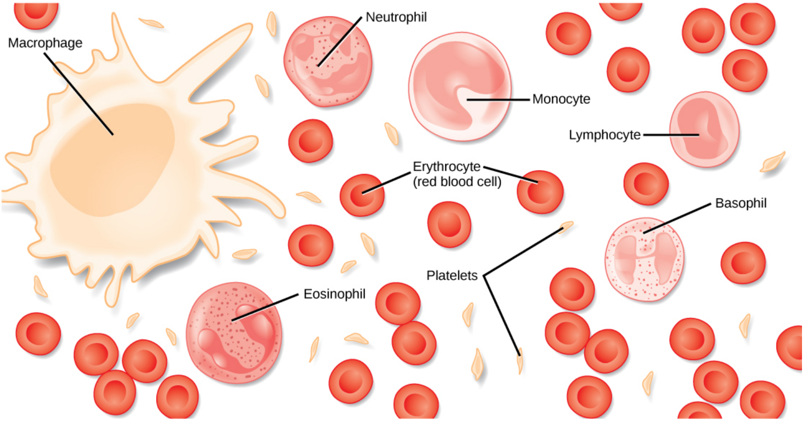
Adipose tissue, or fat tissue, is considered a connective tissue even though it does not have fibroblasts or a real matrix and only has a few fibers. Adipose tissue is made up of cells called adipocytes that collect and store fat in the form of triglycerides, for energy metabolism. Adipose tissues additionally serve as insulation to help maintain body temperatures, allowing animals to be

endothermic, and they function as cushioning against damage to body organs. Under a microscope, adipose tissue cells appear empty due to the extraction of fat during the processing of the material for viewing, as seen in [link]. The thin lines in the image are the cell membranes, and the nuclei are the small, black dots at the edges of the cells.



Blood

Blood is considered a connective tissue because it has a matrix, as shown in [link]. The living cell types are red blood cells (RBC), also called erythrocytes, and white blood cells (WBC), also called leukocytes. The fluid portion of whole blood, its matrix, is commonly called plasma.



The cell found in greatest abundance in blood is the erythrocyte. Erythrocytes are counted in millions in a blood sample: the average number of red blood cells in primates is 4.7 to 5.5 million cells per microliter. Erythrocytes are consistently the same size in a species, but vary in size between species. For example, the average diameter of a primate red blood cell is 7.5 µl, a dog is close at 7.0 µl, but a cat's RBC diameter is 5.9 µl. Sheep erythrocytes are even smaller at 4.6 µl. Mammalian erythrocytes lose their nuclei and mitochondria when they are released from the bone marrow where they are made. Fish, amphibian, and avian red blood cells maintain their nuclei and mitochondria throughout the cell's life. The principal job of an erythrocyte is to carry and deliver oxygen to the tissues.

Leukocytes are the predominant white blood cells found in the peripheral blood. Leukocytes are counted in the thousands in the blood with measurements expressed as ranges: primate counts

range from 4,800 to 10,800 cells per µl, dogs from 5,600 to 19,200 cells per µl, cats from 8,000 to 25,000 cells per µl, cattle from 4,000 to 12,000 cells per µl, and pigs from 11,000 to 22,000 cells per µl.

Lymphocytes function primarily in the immune response to foreign antigens or material. Different types of lymphocytes make antibodies tailored to the foreign antigens and control the production of those antibodies. Neutrophils are phagocytic cells and they participate in one of the early lines of defense against microbial invaders, aiding in the removal of bacteria that has entered the body. Another leukocyte that is found in the peripheral blood is the monocyte. Monocytes give rise to phagocytic macrophages that clean up dead and damaged cells in the body, whether they are foreign or from the host animal. Two additional leukocytes in the blood are eosinophils and basophils—both help to facilitate the inflammatory response.

The slightly granular material among the cells is a cytoplasmic fragment of a cell in the bone marrow. This is called a platelet or thrombocyte. Platelets participate in the stages leading up to coagulation of the blood to stop bleeding through damaged blood vessels. Blood has a number of functions, but primarily it transports material through the body to bring nutrients to cells and remove waste material from them.

Smooth muscle cells do not have striations, while

skeletal muscle cells do. Cardiac muscle cells have striations, but, unlike the multinucleate skeletal cells, they have only one nucleus. Cardiac muscle tissue also has intercalated discs, specialized regions running along the plasma membrane that join adjacent cardiac muscle cells and assist in passing an electrical impulse from cell to cell.

Muscle Tissues

There are three types of muscle in animal bodies: smooth, skeletal, and cardiac. They differ by the presence or absence of striations or bands, the number and location of nuclei, whether they are voluntarily or involuntarily controlled, and their location within the body. [link] summarizes these differences.

Types of Muscles

Type of Muscle	Striations	Nuclei	Control	Location
smooth	no	single, in center	involuntary	visceral organs
skeletal	yes	many, at periphery	voluntary	skeletal muscles

cardiac	yes	single, in center	involuntary	heart

Smooth Muscle

Smooth muscle does not have striations in its cells. It has a single, centrally located nucleus, as shown in [link]. Constriction of smooth muscle occurs under involuntary, autonomic nervous control and in response to local conditions in the tissues. Smooth muscle tissue is also called non-striated as it lacks the banded appearance of skeletal and cardiac muscle. The walls of blood vessels, the tubes of the digestive system, and the tubes of the reproductive systems are composed of mostly smooth muscle.

Skeletal Muscle

Skeletal muscle has striations across its cells caused by the arrangement of the contractile proteins actin and myosin. These muscle cells are relatively long and have multiple nuclei along the edge of the cell. Skeletal muscle is under voluntary, somatic nervous system control and is found in the muscles that move bones. [link] illustrates the histology of skeletal muscle.

Cardiac Muscle

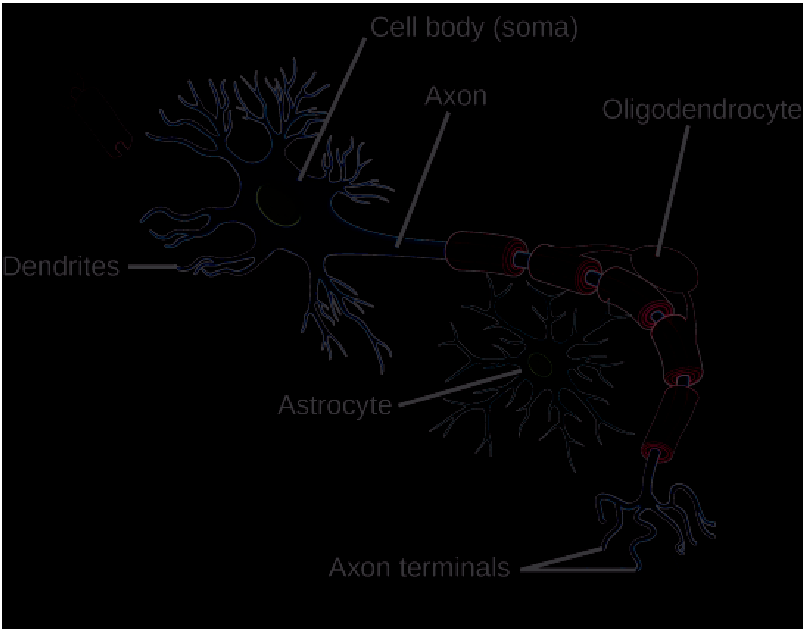
Cardiac muscle, shown in [link], is found only in the heart. Like skeletal muscle, it has cross striations in its cells, but cardiac muscle has a single, centrally located nucleus. Cardiac muscle is not under voluntary control but can be influenced by the autonomic nervous system to speed up or slow down. An added feature to cardiac muscle cells is a line than extends along the end of the cell as it abuts the next cardiac cell in the row. This line is called an intercalated disc: it assists in passing electrical impulse efficiently from one cell to the next and maintains the strong connection between neighboring cardiac cells.

The neuron has projections called dendrites that receive signals and projections called axons that send signals. Also shown are two types of glial cells: astrocytes regulate the chemical environment of the nerve cell, and oligodendrocytes insulate the axon so the electrical nerve impulse is transferred more efficiently.

Nervous Tissues

Nervous tissues are made of cells specialized to receive and transmit electrical impulses from specific areas of the body and to send them to specific locations in the body. The main cell of the nervous system is the neuron, illustrated in [link]. The large structure with a central nucleus is the cell

body of the neuron. Projections from the cell body are either dendrites specialized in receiving input or a single axon specialized in transmitting impulses. Some glial cells are also shown. Astrocytes regulate the chemical environment of the nerve cell, and oligodendrocytes insulate the axon so the electrical nerve impulse is transferred more efficiently. Other glial cells that are not shown support the nutritional and waste requirements of the neuron. Some of the glial cells are phagocytic and remove debris or damaged cells from the tissue. A nerve consists of neurons and glial cells.



Link to Learning
Click through the interactive review to learn more

about epithelial tissues.

Career Connections

Pathologist

A pathologist is a medical doctor or veterinarian who has specialized in the laboratory detection of disease in animals, including humans. These professionals complete medical school education and follow it with an extensive post-graduate residency at a medical center. A pathologist may oversee clinical laboratories for the evaluation of body tissue and blood samples for the detection of disease or infection. They examine tissue specimens through a microscope to identify cancers and other diseases. Some pathologists perform autopsies to determine the cause of death and the progression of disease.

Section Summary

The basic building blocks of complex animals are four primary tissues. These are combined to form organs, which have a specific, specialized function within the body, such as the skin or kidney. Organs

are organized together to perform common functions in the form of systems. The four primary tissues are epithelia, connective tissues, muscle tissues, and nervous tissues.

Visual Connection Questions

[link] Which of the following statements about types of epithelial cells is false?

1. Simple columnar epithelial cells line the tissue of the lung.
2. Simple cuboidal epithelial cells are involved in the filtering of blood in the kidney.
3. Pseudostratisfied columnar epithilia occur in a single layer, but the arrangement of nuclei makes it appear that more than one layer is present.
4. Transitional epithelia change in thickness depending on how full the bladder is.

[link] A

Review Questions

Which type of epithelial cell is best adapted to aid diffusion?

1. squamous
2. cuboidal
3. columnar
4. transitional

C

Which type of epithelial cell is found in glands?

1. squamous
2. cuboidal
3. columnar
4. transitional

B

Which type of epithelial cell is found in the urinary bladder?

1. squamous
2. cuboidal
3. columnar
4. transitional

D

Which type of connective tissue has the most fibers?

1. loose connective tissue
2. fibrous connective tissue
3. cartilage
4. bone

B

Which type of connective tissue has a mineralized different matrix?

1. loose connective tissue
2. fibrous connective tissue
3. cartilage
4. bone

D

The cell found in bone that breaks it down is called an ______.

1. osteoblast
2. osteocyte

3. osteoclast
4. osteon

C

The cell found in bone that makes the bone is called an ______.

1. osteoblast
2. osteocyte
3. osteoclast
4. osteon

A

Plasma is the ______.

1. fibers in blood
2. matrix of blood
3. cell that phagocytizes bacteria
4. cell fragment found in the tissue

B

The type of muscle cell under voluntary control is the ______.

1. smooth muscle
2. skeletal muscle
3. cardiac muscle
4. visceral muscle

B

The part of a neuron that contains the nucleus is the

1. cell body
2. dendrite
3. axon
4. glial

A

Why are intercalated discs essential to the function of cardiac muscle?

1. The discs maintain the barriers between the cells.
2. The discs pass nutrients between cells.
3. The discs ensure that all the cardiac muscle cells beat as a single unit.
4. The discs control the heart rate.

C

Critical Thinking Questions

How can squamous epithelia both facilitate diffusion and prevent damage from abrasion?

Squamous epithelia can be either simple or stratified. As a single layer of cells, it presents a very thin epithelia that minimally inhibits diffusion. As a stratified epithelia, the surface cells can be sloughed off and the cells in deeper layers protect the underlying tissues from damage.

What are the similarities between cartilage and bone?

Both contain cells other than the traditional fibroblast. Both have cells that lodge in spaces within the tissue called lacunae. Both collagen and elastic fibers are found in bone and cartilage. Both tissues participate in vertebrate skeletal development and formation.

Multiple sclerosis is a debilitating autoimmune disease that results in the loss of the insulation around neuron axons. What cell type is the immune system attacking, and how does this disrupt the transfer of messages by the nervous system?

In multiple sclerosis, the immune system attacks the oligodendrocytes. The death of oligodendrocytes results in the loss of the insulating sheath around the axon of the neurons. When the sheath is gone, the electrical impulses travel much more slowly down the length of the axon.

When a person leads a sedentary life his skeletal muscles atrophy, but his smooth muscles do not. Why?

Skeletal muscles are involved in voluntary motion, so the person has to make the choice to work those muscles through exercise or movement. Smooth muscles are involved in involuntary activities of the body (ex. blood vessel expansion and contraction, intestinal peristalsis) so they are active even when a person is sedentary.

Glossary

canaliculus
: microchannel that connects the lacunae and aids diffusion between cells

cartilage
: type of connective tissue with a large amount of ground substance matrix, cells called chondrocytes, and some amount of fibers

chondrocyte
: cell found in cartilage

columnar epithelia
: epithelia made of cells taller than they are wide, specialized in absorption

connective tissue
: type of tissue made of cells, ground substance matrix, and fibers

cuboidal epithelia
: epithelia made of cube-shaped cells, specialized in glandular functions

epithelial tissue
: tissue that either lines or covers organs or other tissues

fibrous connective tissue
: type of connective tissue with a high

concentration of fibers

lacuna
space in cartilage and bone that contains living cells

loose (areolar) connective tissue
type of connective tissue with small amounts of cells, matrix, and fibers; found around blood vessels

matrix
component of connective tissue made of both living and nonliving (ground substances) cells

osteon
subunit of compact bone

pseudostratified
layer of epithelia that appears multilayered, but is a simple covering

simple epithelia
single layer of epithelial cells

squamous epithelia
type of epithelia made of flat cells, specialized in aiding diffusion or preventing abrasion

stratified epithelia
multiple layers of epithelial cells

trabecula

tiny plate that makes up spongy bone and gives it strength

transitional epithelia
: epithelia that can transition for appearing multilayered to simple; also called uroepithelial

Homeostasis
By the end of this section, you will be able to do the following:

- Define homeostasis
- Describe the factors affecting homeostasis
- Discuss positive and negative feedback mechanisms used in homeostasis
- Describe thermoregulation of endothermic and ectothermic animals

Animal organs and organ systems constantly adjust to internal and external changes through a process called homeostasis ("steady state"). These changes might be in the level of glucose or calcium in blood or in external temperatures. **Homeostasis** means to maintain dynamic equilibrium in the body. It is dynamic because it is constantly adjusting to the changes that the body's systems encounter. It is equilibrium because body functions are kept within specific ranges. Even an animal that is apparently inactive is maintaining this homeostatic equilibrium.

Homeostatic Process

The goal of homeostasis is the maintenance of equilibrium around a point or value called a **set point**. While there are normal fluctuations from the

set point, the body's systems will usually attempt to go back to this point. A change in the internal or external environment is called a stimulus and is detected by a receptor; the response of the system is to adjust the deviation parameter toward the set point. For instance, if the body becomes too warm, adjustments are made to cool the animal. If the blood's glucose rises after a meal, adjustments are made to lower the blood glucose level by getting the nutrient into tissues that need it or to store it for later use.

Blood sugar levels are controlled by a negative feedback loop. (credit: modification of work by Jon Sullivan)

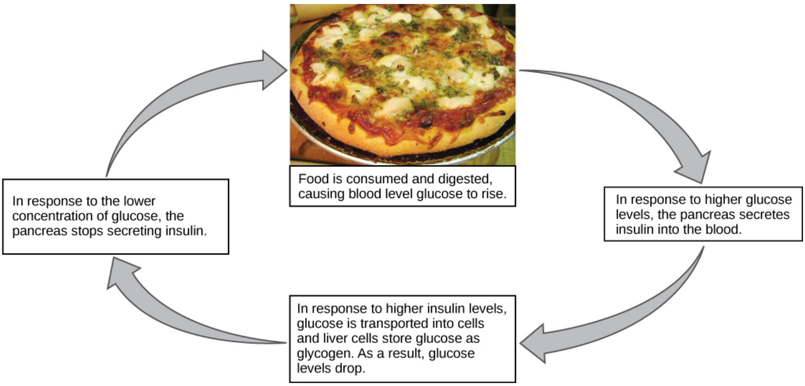
Control of Homeostasis

When a change occurs in an animal's environment, an adjustment must be made. The receptor senses the change in the environment, then sends a signal to the control center (in most cases, the brain) which in turn generates a response that is signaled to an effector. The effector is a muscle (that contracts or relaxes) or a gland that secretes. Homeostatsis is maintained by negative feedback loops. Positive feedback loops actually push the organism further out of homeostasis, but may be necessary for life to occur. Homeostasis is controlled by the nervous and endocrine system of mammals.

Negative Feedback Mechanisms

Any homeostatic process that changes the direction of the stimulus is a **negative feedback loop**. It may either increase or decrease the stimulus, but the stimulus is not allowed to continue as it did before the receptor sensed it. In other words, if a level is too high, the body does something to bring it down, and conversely, if a level is too low, the body does something to make it go up. Hence the term negative feedback. An example is animal maintenance of blood glucose levels. When an animal has eaten, blood glucose levels rise. This is sensed by the nervous system. Specialized cells in the pancreas sense this, and the hormone insulin is released by the endocrine system. Insulin causes blood glucose levels to decrease, as would be expected in a negative feedback system, as illustrated in [link]. However, if an animal has not eaten and blood glucose levels decrease, this is sensed in another group of cells in the pancreas, and the hormone glucagon is released causing glucose levels to increase. This is still a negative feedback loop, but not in the direction expected by the use of the term "negative." Another example of an increase as a result of the feedback loop is the control of blood calcium. If calcium levels decrease, specialized cells in the parathyroid gland sense this and release parathyroid hormone (PTH), causing an increased absorption of calcium through the intestines and kidneys and, possibly, the breakdown

of bone in order to liberate calcium. The effects of PTH are to raise blood levels of the element. Negative feedback loops are the predominant mechanism used in homeostasis.



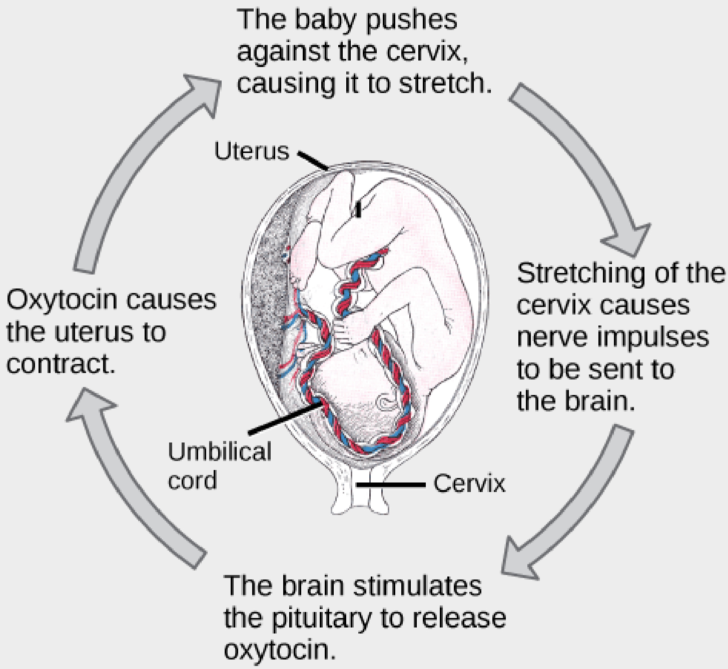
Positive Feedback Loop

A **positive feedback loop** maintains the direction of the stimulus, possibly accelerating it. Few examples of positive feedback loops exist in animal bodies, but one is found in the cascade of chemical reactions that result in blood clotting, or coagulation. As one clotting factor is activated, it activates the next factor in sequence until a fibrin clot is achieved. The direction is maintained, not changed, so this is positive feedback. Another example of positive feedback is uterine contractions during childbirth, as illustrated in [link]. The hormone oxytocin, made by the endocrine system, stimulates the contraction of the uterus. This produces pain sensed by the nervous system. Instead of lowering the oxytocin and causing the pain to

subside, more oxytocin is produced until the contractions are powerful enough to produce childbirth.

Visual Connection

The birth of a human infant is the result of positive feedback.



State whether each of the following processes is regulated by a positive feedback loop or a negative feedback loop.

1. A person feels satiated after eating a large meal.
2. The blood has plenty of red blood cells. As a

result, erythropoietin, a hormone that stimulates the production of new red blood cells, is no longer released from the kidney.

Set Point

It is possible to adjust a system's set point. When this happens, the feedback loop works to maintain the new setting. An example of this is blood pressure: over time, the normal or set point for blood pressure can increase as a result of continued increases in blood pressure. The body no longer recognizes the elevation as abnormal and no attempt is made to return to the lower set point. The result is the maintenance of an elevated blood pressure that can have harmful effects on the body. Medication can lower blood pressure and lower the set point in the system to a more healthy level. This is called a process of **alteration** of the set point in a feedback loop.

Changes can be made in a group of body organ systems in order to maintain a set point in another system. This is called **acclimatization**. This occurs, for instance, when an animal migrates to a higher altitude than that to which it is accustomed. In order to adjust to the lower oxygen levels at the new altitude, the body increases the number of red blood cells circulating in the blood to ensure adequate

oxygen delivery to the tissues. Another example of acclimatization is animals that have seasonal changes in their coats: a heavier coat in the winter ensures adequate heat retention, and a light coat in summer assists in keeping body temperature from rising to harmful levels.

Link to Learning

Feedback mechanisms can be understood in terms of driving a race car along a track: watch a short video lesson on positive and negative feedback loops.

https://www.openstax.org/l/feedback_loops

Homeostasis: Thermoregulation

Body temperature affects body activities. Generally, as body temperature rises, enzyme activity rises as well. For every ten degree centigrade rise in temperature, enzyme activity doubles, up to a point. Body proteins, including enzymes, begin to denature and lose their function with high heat (around 50$_o$C for mammals). Enzyme activity will decrease by half for every ten degree centigrade drop in temperature, to the point of freezing, with a few exceptions. Some

fish can withstand freezing solid and return to normal with thawing.

Link to Learning
Watch this Discovery Channel video on thermoregulation to see illustrations of this process in a variety of animals.
https://www.openstax.org/l/thermoregulate

Heat can be exchanged by four mechanisms: (a) radiation, (b) evaporation, (c) convection, or (d) conduction. (credit b: modification of work by "Kullez"/Flickr; credit c: modification of work by Chad Rosenthal; credit d: modification of work by "stacey.d"/Flickr)

Endotherms and Ectotherms

Animals can be divided into two groups: some maintain a constant body temperature in the face of differing environmental temperatures, while others have a body temperature that is the same as their environment and thus varies with the environment. Animals that rely on external temperatures to set their body temperature are ectotherms. This group has been called cold-blooded, but the term may not apply to an animal in the desert with a very warm

body temperature. In contrast to ectotherms, poikilotherms are animals with constantly varying internal temperatures. An animal that maintains a constant body temperature in the face of environmental changes is called a homeotherm. Endotherms are animals that rely on internal sources for maintenance of relatively constant body temperature in varying environmental temperatures. These animals are able to maintain a level of metabolic activity at cooler temperature, which an ectotherm cannot due to differing enzyme levels of activity. It is worth mentioning that some ectotherms and poikilotherms have relatively constant body temperatures due to the constant environmental temperatures in their habitats. These animals are so-called ectothermic homeotherms, like some deep sea fish species.

Heat can be exchanged between an animal and its environment through four mechanisms: radiation, evaporation, convection, and conduction ([link]). Radiation is the emission of electromagnetic "heat" waves. Heat comes from the sun in this manner and radiates from dry skin the same way. Heat can be removed with liquid from a surface during evaporation. This occurs when a mammal sweats. Convection currents of air remove heat from the surface of dry skin as the air passes over it. Heat will be conducted from one surface to another during direct contact with the surfaces, such as an animal resting on a warm rock.

(a) Radiation (b) Evaporation

(c) Convection (d) Conduction

Heat Conservation and Dissipation

Animals conserve or dissipate heat in a variety of ways. In certain climates, endothermic animals have some form of insulation, such as fur, fat, feathers, or some combination thereof. Animals with thick fur or feathers create an insulating layer of air between their skin and internal organs. Polar bears and seals live and swim in a subfreezing environment and yet maintain a constant, warm, body temperature. The arctic fox, for example, uses its fluffy tail as extra insulation when it curls up to sleep in cold weather. Mammals have a residual effect from shivering and increased muscle activity: arrector pili muscles cause “goose bumps,” causing small hairs to stand

up when the individual is cold; this has the intended effect of increasing body temperature. Mammals use layers of fat to achieve the same end. Loss of significant amounts of body fat will compromise an individual's ability to conserve heat.

Endotherms use their circulatory systems to help maintain body temperature. Vasodilation brings more blood and heat to the body surface, facilitating radiation and evaporative heat loss, which helps to cool the body. Vasoconstriction reduces blood flow in peripheral blood vessels, forcing blood toward the core and the vital organs found there, and conserving heat. Some animals have adaptions to their circulatory system that enable them to transfer heat from arteries to veins, warming blood returning to the heart. This is called a countercurrent heat exchange; it prevents the cold venous blood from cooling the heart and other internal organs. This adaption can be shut down in some animals to prevent overheating the internal organs. The countercurrent adaption is found in many animals, including dolphins, sharks, bony fish, bees, and hummingbirds. In contrast, similar adaptations can help cool endotherms when needed, such as dolphin flukes and elephant ears.

Some ectothermic animals use changes in their behavior to help regulate body temperature. For example, a desert ectothermic animal may simply seek cooler areas during the hottest part of the day

in the desert to keep from getting too warm. The same animals may climb onto rocks to capture heat during a cold desert night. Some animals seek water to aid evaporation in cooling them, as seen with reptiles. Other ectotherms use group activity such as the activity of bees to warm a hive to survive winter.

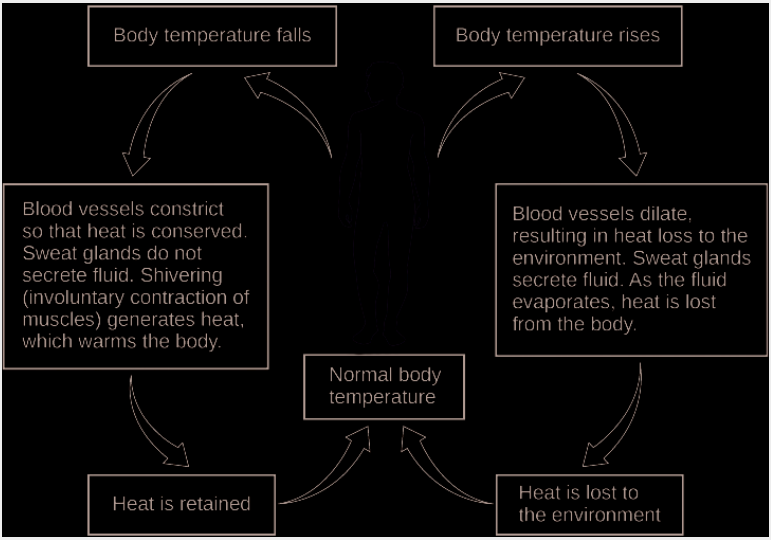
Many animals, especially mammals, use metabolic waste heat as a heat source. When muscles are contracted, most of the energy from the ATP used in muscle actions is wasted energy that translates into heat. Severe cold elicits a shivering reflex that generates heat for the body. Many species also have a type of adipose tissue called brown fat that specializes in generating heat.

Neural Control of Thermoregulation

The nervous system is important to **thermoregulation**, as illustrated in [link]. The processes of homeostasis and temperature control are centered in the hypothalamus of the advanced animal brain.

Visual Connection

The body is able to regulate temperature in response to signals from the nervous system.



When bacteria are destroyed by leuckocytes, pyrogens are released into the blood. Pyrogens reset the body’s thermostat to a higher temperature, resulting in fever. How might pyrogens cause the body temperature to rise?

The hypothalamus maintains the set point for body temperature through reflexes that cause vasodilation and sweating when the body is too warm, or vasoconstriction and shivering when the body is too cold. It responds to chemicals from the body. When a bacterium is destroyed by phagocytic leukocytes, chemicals called endogenous pyrogens are released into the blood. These pyrogens circulate to the hypothalamus and reset the thermostat. This allows the body’s temperature to increase in what is commonly called a fever. An increase in body

temperature causes iron to be conserved, which reduces a nutrient needed by bacteria. An increase in body heat also increases the activity of the animal's enzymes and protective cells while inhibiting the enzymes and activity of the invading microorganisms. Finally, heat itself may also kill the pathogen. A fever that was once thought to be a complication of an infection is now understood to be a normal defense mechanism.

Section Summary

Homeostasis is a dynamic equilibrium that is maintained in body tissues and organs. It is dynamic because it is constantly adjusting to the changes that the systems encounter. It is in equilibrium because body functions are kept within a normal range, with some fluctuations around a set point for the processes.

Visual Connection Questions

[link] State whether each of the following processes are regulated by a positive feedback loop or a negative feedback loop.

1. A person feels satiated after eating a large

meal.

2. The blood has plenty of red blood cells. As a result, erythropoietin, a hormone that stimulates the production of new red blood cells, is no longer released from the kidney.

[link] Both processes are the result of negative feedback loops. Negative feedback loops, which tend to keep a system at equilibrium, are more common than positive feedback loops.

[link] When bacteria are destroyed by leuckocytes, pyrogens are released into the blood. Pyrogens reset the body's thermostat to a higher temperature, resulting in fever. How might pyrogens cause the body temperature to rise?

[link] Pyrogens increase body temperature by causing the blood vessels to constrict, inducing shivering, and stopping sweat glands from secreting fluid.

Review Questions

When faced with a sudden drop in environmental temperature, an endothermic animal will:

1. experience a drop in its body temperature
2. wait to see if it goes lower
3. increase muscle activity to generate heat
4. add fur or fat to increase insulation

C

Which is an example of negative feedback?

1. lowering of blood glucose after a meal
2. blood clotting after an injury
3. lactation during nursing
4. uterine contractions during labor

A

Which method of heat exchange occurs during direct contact between the source and animal?

1. radiation
2. evaporation
3. convection
4. conduction

D

The body's thermostat is located in the ______.

1. homeostatic receptor
2. hypothalamus
3. medulla
4. vasodilation center

B

Which of the following is **not** true about acclimatization?

1. Acclimatization allows animals to compensate for changes in their environment.
2. Acclimatization improves function in a new environment.
3. Acclimatization occurs when an animal tries to reestablish a homeostatic set point.
4. Acclimatization is passed on to offspring of acclimated individuals.

D

Which of the following is **not** a way that

ectotherms can change their body temperatures?

1. Sweating for evaporative cooling.
2. Adjusting the timing of their daily activities.
3. Seek out or avoid direct sunlight.
4. Huddle in a group.

A

Critical Thinking Questions

Why are negative feedback loops used to control body homeostasis?

An adjustment to a change in the internal or external environment requires a change in the direction of the stimulus. A negative feedback loop accomplishes this, while a positive feedback loop would continue the stimulus and result in harm to the animal.

Why is a fever a “good thing” during a bacterial infection?

Mammalian enzymes increase activity to the point of denaturation, increasing the chemical activity of the cells involved. Bacterial enzymes have a specific temperature for their most efficient activity and are inhibited at either higher or lower temperatures. Fever results in an increase in the destruction of the invading bacteria by increasing the effectiveness of body defenses and an inhibiting bacterial metabolism.

How is a condition such as diabetes a good example of the failure of a set point in humans?

Diabetes is often associated with a lack in production of insulin. Without insulin, blood glucose levels go up after a meal, but never go back down to normal levels.

On a molecular level, how can endotherms produce their own heat by adjusting processes associated with cellular respiration? If needed, review Ch. 7 for details on respiration.

Animals are capable of thermal uncoupling when they need to generate heat to maintain their body temperatures. In this process, an

uncoupling protein provides a channel in the inner mitochondrial membrane that allows protons to leave the lumen without moving through the ATP synthase. This generates heat rather than chemical energy as the final product of cellular respiration.

Glossary

acclimatization
: alteration in a body system in response to environmental change

alteration
: change of the set point in a homeostatic system

homeostasis
: dynamic equilibrium maintaining appropriate body functions

negative feedback loop
: feedback to a control mechanism that increases or decreases a stimulus instead of maintaining it

positive feedback loop
: feedback to a control mechanism that continues the direction of a stimulus

set point

midpoint or target point in homeostasis

thermoregulation
regulation of body temperature

Introduction

class="introduction" For humans, fruits and vegetables are important in maintaining a balanced diet. (credit: modification of work by Julie Rybarczyk)

All living organisms need nutrients to survive. While plants can obtain the molecules required for cellular function through the process of photosynthesis, most animals obtain their nutrients by the consumption of other organisms. At the cellular level, the biological molecules necessary for animal function are amino acids, lipid molecules, nucleotides, and simple sugars. However, the food consumed consists of protein, fat, and complex carbohydrates. Animals must convert these macromolecules into the simple molecules required for maintaining cellular functions, such as assembling new molecules, cells, and tissues. The

conversion of the food consumed to the nutrients required is a multistep process involving digestion and absorption. During digestion, food particles are broken down to smaller components, and later, they are absorbed by the body.

One of the challenges in human nutrition is maintaining a balance between food intake, storage, and energy expenditure. Imbalances can have serious health consequences. For example, eating too much food while not expending much energy leads to obesity, which in turn will increase the risk of developing illnesses such as type-2 diabetes and cardiovascular disease. The recent rise in obesity and related diseases makes understanding the role of diet and nutrition in maintaining good health all the more important.

Digestive Systems

By the end of this section, you will be able to do the following:

- Explain the processes of digestion and absorption
- Compare and contrast different types of digestive systems
- Explain the specialized functions of the organs involved in processing food in the body
- Describe the ways in which organs work together to digest food and absorb nutrients

Animals obtain their nutrition from the consumption of other organisms. Depending on their diet, animals can be classified into the following categories: plant eaters (herbivores), meat eaters (carnivores), and those that eat both plants and animals (omnivores). The nutrients and macromolecules present in food are not immediately accessible to the cells. There are a number of processes that modify food within the animal body in order to make the nutrients and organic molecules accessible for cellular function. As animals evolved in complexity of form and function, their digestive systems have also evolved to accommodate their various dietary needs.

Herbivores, like this (a) mule deer and (b) monarch caterpillar, eat primarily plant material. (credit a: modification of work by Bill Ebbesen; credit b: modification of work by Doug Bowman) Carnivores like the (a) lion eat primarily meat. The (b) ladybug

is also a carnivore that consumes small insects called aphids. (credit a: modification of work by Kevin Pluck; credit b: modification of work by Jon Sullivan) Omnivores like the (a) bear and (b) crayfish eat both plant and animal based food. (credit a: modification of work by Dave Menke; credit b: modification of work by Jon Sullivan)

Herbivores, Omnivores, and Carnivores

Herbivores are animals whose primary food source is plant-based. Examples of herbivores, as shown in [link] include vertebrates like deer, koalas, and some bird species, as well as invertebrates such as crickets and caterpillars. These animals have evolved digestive systems capable of handling large amounts of plant material. Herbivores can be further classified into frugivores (fruit-eaters), granivores (seed eaters), nectivores (nectar feeders), and folivores (leaf eaters).


(a) (b)

Carnivores are animals that eat other animals. The word carnivore is derived from Latin and literally means "meat eater." Wild cats such as lions, shown in [link]**a** and tigers are examples of vertebrate carnivores, as are snakes and sharks, while invertebrate carnivores include sea stars, spiders, and ladybugs, shown in [link]**b**. Obligate carnivores are those that rely entirely on animal flesh to obtain their nutrients; examples of obligate carnivores are members of the cat family, such as lions and cheetahs. Facultative carnivores are those that also eat non-animal food in addition to animal food. Note that there is no clear line that differentiates facultative carnivores from omnivores; dogs would be considered facultative carnivores.

(a) (b)

Omnivores are animals that eat both plant- and animal-derived food. In Latin, omnivore means to eat everything. Humans, bears (shown in [link]**a**), and chickens are example of vertebrate omnivores; invertebrate omnivores include cockroaches and crayfish (shown in [link]**b**).

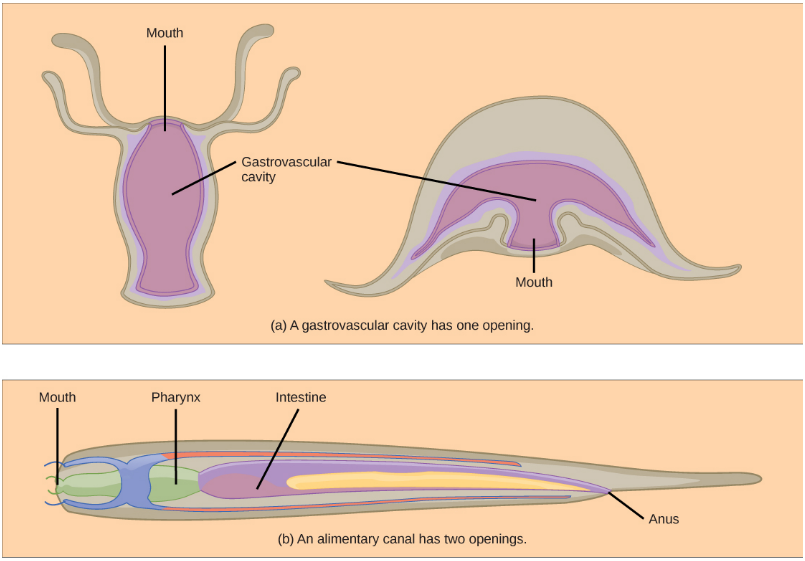
(a) (b)

(a) A gastrovascular cavity has a single opening through which food is ingested and waste is excreted, as shown in this hydra and in this jellyfish medusa. (b) An alimentary canal has two openings: a mouth for ingesting food, and an anus for eliminating waste, as shown in this nematode.

Invertebrate Digestive Systems

Animals have evolved different types of digestive systems to aid in the digestion of the different foods they consume. The simplest example is that of a **gastrovascular cavity** and is found in organisms with only one opening for digestion. Platyhelminthes (flatworms), Ctenophora (comb jellies), and Cnidaria (coral, jelly fish, and sea anemones) use this type of digestion. Gastrovascular cavities, as shown in [link]**a**, are typically a blind tube or cavity with only one opening, the “mouth”, which also serves as an “anus”. Ingested material enters the mouth and passes through a hollow, tubular cavity. Cells within the cavity secrete digestive enzymes that breakdown the food. The food particles are engulfed by the cells lining the gastrovascular cavity.

The **alimentary canal**, shown in [link]**b**, is a more advanced system: it consists of one tube with a mouth at one end and an anus at the other. Earthworms are an example of an animal with an alimentary canal. Once the food is ingested through the mouth, it passes through the esophagus and is stored in an organ called the crop; then it passes into the gizzard where it is churned and digested. From the gizzard, the food passes through the intestine, the nutrients are absorbed, and the waste is eliminated as feces, called castings, through the anus.



(a) Humans and herbivores, such as the (b) rabbit, have a monogastric digestive system. However, in the rabbit the small intestine and cecum are enlarged to allow more time to digest plant material. The enlarged organ provides more surface area for absorption of nutrients. Rabbits digest their food twice: the first time food passes through the digestive system, it collects in the cecum, and then it passes as soft feces called cecotrophes. The rabbit re-ingests these cecotrophes to further digest them. The avian esophagus has a pouch, called a crop, which stores food. Food passes from the crop to the first of two stomachs, called the proventriculus, which contains digestive juices that breakdown food. From the proventriculus, the food enters the second stomach, called the gizzard, which grinds food. Some birds swallow stones or grit, which are stored in the gizzard, to aid the grinding process. Birds do not have separate openings to excrete urine

and feces. Instead, uric acid from the kidneys is secreted into the large intestine and combined with waste from the digestive process. This waste is excreted through an opening called the cloaca. Ruminant animals, such as goats and cows, have four stomachs. The first two stomachs, the rumen and the reticulum, contain prokaryotes and protists that are able to digest cellulose fiber. The ruminant regurgitates cud from the reticulum, chews it, and swallows it into a third stomach, the omasum, which removes water. The cud then passes onto the fourth stomach, the abomasum, where it is digested by enzymes produced by the ruminant.

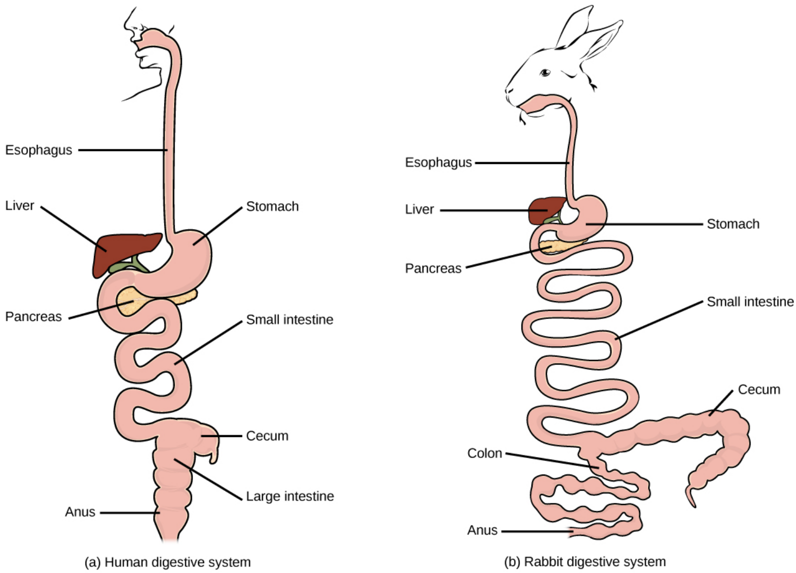
Vertebrate Digestive Systems

Vertebrates have evolved more complex digestive systems to adapt to their dietary needs. Some animals have a single stomach, while others have multi-chambered stomachs. Birds have developed a digestive system adapted to eating unmasticated food.

Monogastric: Single-chambered Stomach

As the word **monogastric** suggests, this type of digestive system consists of one ("mono") stomach chamber ("gastric"). Humans and many animals have a monogastric digestive system as illustrated in [link]**ab**. The process of digestion begins with the

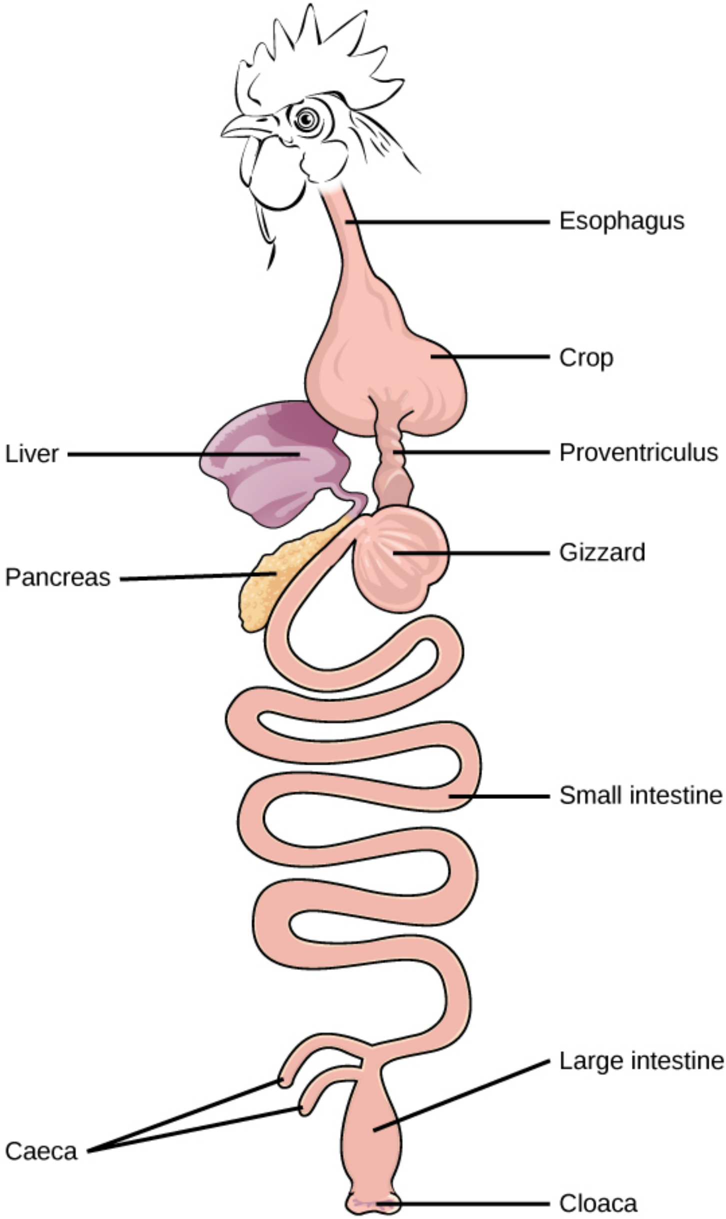
mouth and the intake of food. The teeth play an important role in masticating (chewing) or physically breaking down food into smaller particles. The enzymes present in saliva also begin to chemically breakdown food. The esophagus is a long tube that connects the mouth to the stomach. Using peristalsis, or wave-like smooth muscle contractions, the muscles of the esophagus push the food towards the stomach. In order to speed up the actions of enzymes in the stomach, the stomach is an extremely acidic environment, with a pH between 1.5 and 2.5. The gastric juices, which include enzymes in the stomach, act on the food particles and continue the process of digestion. Further breakdown of food takes place in the small intestine where enzymes produced by the liver, the small intestine, and the pancreas continue the process of digestion. The nutrients are absorbed into the bloodstream across the epithelial cells lining the walls of the small intestines. The waste material travels on to the large intestine where water is absorbed and the drier waste material is compacted into feces; it is stored until it is excreted through the rectum.



Avian

Birds face special challenges when it comes to obtaining nutrition from food. They do not have teeth and so their digestive system, shown in [link], must be able to process un-masticated food. Birds have evolved a variety of beak types that reflect the vast variety in their diet, ranging from seeds and insects to fruits and nuts. Because most birds fly, their metabolic rates are high in order to efficiently process food and keep their body weight low. The stomach of birds has two chambers: the **proventriculus**, where gastric juices are produced to digest the food before it enters the stomach, and the **gizzard**, where the food is stored, soaked, and mechanically ground. The undigested material forms food pellets that are sometimes regurgitated.

Most of the chemical digestion and absorption happens in the intestine and the waste is excreted through the cloaca.



Evolution Connection

Avian Adaptations

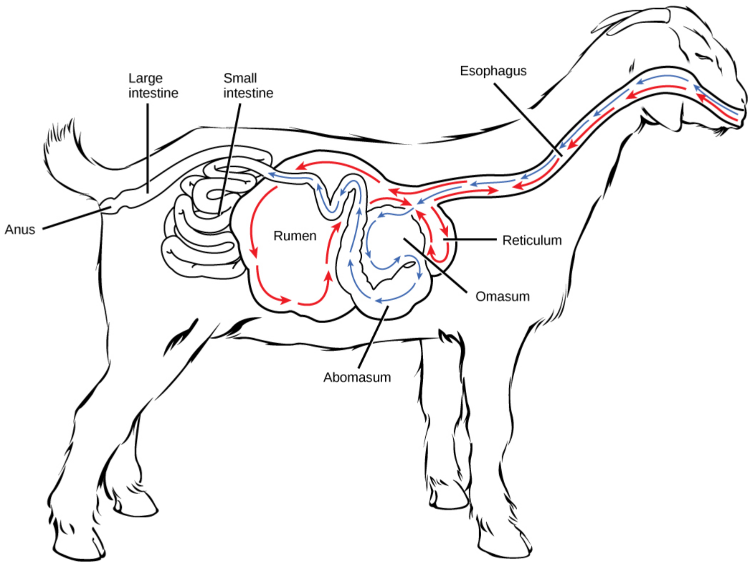
Birds have a highly efficient, simplified digestive system. Recent fossil evidence has shown that the evolutionary divergence of birds from other land animals was characterized by streamlining and simplifying the digestive system. Unlike many other animals, birds do not have teeth to chew their food. In place of lips, they have sharp pointy beaks. The horny beak, lack of jaws, and the smaller tongue of the birds can be traced back to their dinosaur ancestors. The emergence of these changes seems to coincide with the inclusion of seeds in the bird diet. Seed-eating birds have beaks that are shaped for grabbing seeds and the two-compartment stomach allows for delegation of tasks. Since birds need to remain light in order to fly, their metabolic rates are very high, which means they digest their food very quickly and need to eat often. Contrast this with the ruminants, where the digestion of plant matter takes a very long time.

Ruminants

Ruminants are mainly herbivores like cows, sheep, and goats, whose entire diet consists of eating large amounts of **roughage** or fiber. They have evolved digestive systems that help them digest vast

amounts of cellulose. An interesting feature of the ruminants' mouth is that they do not have upper incisor teeth. They use their lower teeth, tongue and lips to tear and chew their food. From the mouth, the food travels to the esophagus and on to the stomach.

To help digest the large amount of plant material, the stomach of the ruminants is a multi-chambered organ, as illustrated in [link]. The four compartments of the stomach are called the rumen, reticulum, omasum, and abomasum. These chambers contain many microbes that breakdown cellulose and ferment ingested food. The abomasum is the "true" stomach and is the equivalent of the monogastric stomach chamber where gastric juices are secreted. The four-compartment gastric chamber provides larger space and the microbial support necessary to digest plant material in ruminants. The fermentation process produces large amounts of gas in the stomach chamber, which must be eliminated. As in other animals, the small intestine plays an important role in nutrient absorption, and the large intestine helps in the elimination of waste.



Pseudo-ruminants

Some animals, such as camels and alpacas, are pseudo-ruminants. They eat a lot of plant material and roughage. Digesting plant material is not easy because plant cell walls contain the polymeric sugar molecule cellulose. The digestive enzymes of these animals cannot breakdown cellulose, but microorganisms present in the digestive system can. Therefore, the digestive system must be able to handle large amounts of roughage and breakdown the cellulose. Pseudo-ruminants have a three-chamber stomach in the digestive system. However, their cecum—a pouched organ at the beginning of the large intestine containing many microorganisms that are necessary for the digestion of plant materials—is large and is the site where the

roughage is fermented and digested. These animals do not have a rumen but have an omasum, abomasum, and reticulum.

Digestion of food begins in the (a) oral cavity. Food is masticated by teeth and moistened by saliva secreted from the (b) salivary glands. Enzymes in the saliva begin to digest starches and fats. With the help of the tongue, the resulting bolus is moved into the esophagus by swallowing. (credit: modification of work by the National Cancer Institute) The esophagus transfers food from the mouth to the stomach through peristaltic movements. The large intestine reabsorbs water from undigested food and stores waste material until it is eliminated.

Parts of the Digestive System

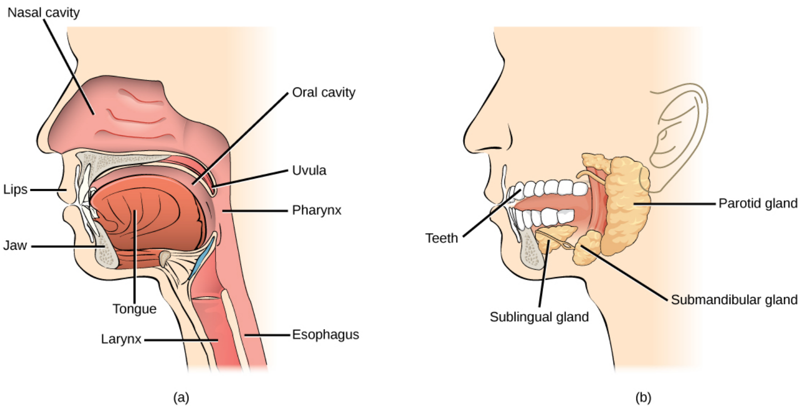
The vertebrate digestive system is designed to facilitate the transformation of food matter into the nutrient components that sustain organisms.

Oral Cavity

The oral cavity, or mouth, is the point of entry of food into the digestive system, illustrated in [link]. The food consumed is broken into smaller particles by mastication, the chewing action of the teeth. All mammals have teeth and can chew their food.

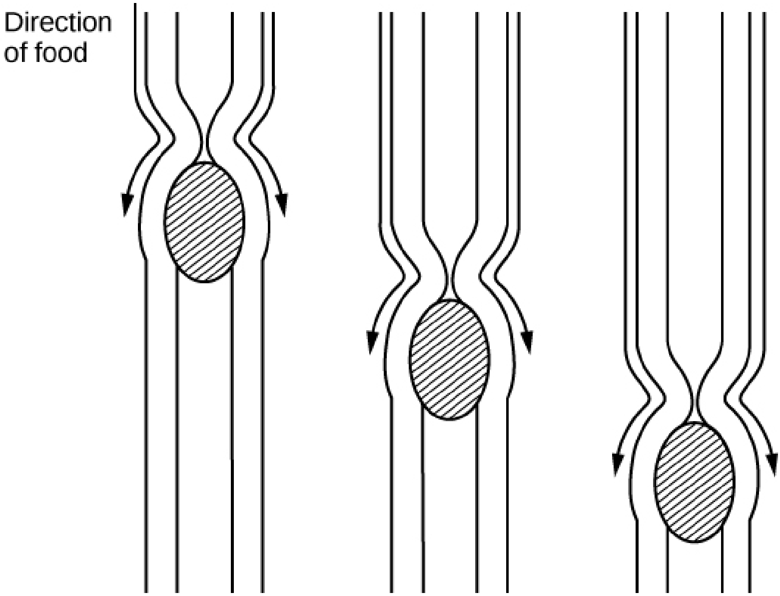
The extensive chemical process of digestion begins

in the mouth. As food is being chewed, saliva, produced by the salivary glands, mixes with the food. Saliva is a watery substance produced in the mouths of many animals. There are three major glands that secrete saliva—the parotid, the submandibular, and the sublingual. Saliva contains mucus that moistens food and buffers the pH of the food. Saliva also contains immunoglobulins and lysozymes, which have antibacterial action to reduce tooth decay by inhibiting growth of some bacteria. Saliva also contains an enzyme called **salivary amylase** that begins the process of converting starches in the food into a disaccharide called maltose. Another enzyme called **lipase** is produced by the cells in the tongue. Lipases are a class of enzymes that can breakdown triglycerides. The lingual lipase begins the breakdown of fat components in the food. The chewing and wetting action provided by the teeth and saliva prepare the food into a mass called the **bolus** for swallowing. The tongue helps in swallowing—moving the bolus from the mouth into the pharynx. The pharynx opens to two passageways: the trachea, which leads to the lungs, and the esophagus, which leads to the stomach. The trachea has an opening called the glottis, which is covered by a cartilaginous flap called the epiglottis. When swallowing, the epiglottis closes the glottis and food passes into the esophagus and not the trachea. This arrangement allows food to be kept out of the trachea.



Esophagus

The **esophagus** is a tubular organ that connects the mouth to the stomach. The chewed and softened food passes through the esophagus after being swallowed. The smooth muscles of the esophagus undergo a series of wave like movements called **peristalsis** that push the food toward the stomach, as illustrated in [link]. The peristalsis wave is unidirectional—it moves food from the mouth to the stomach, and reverse movement is not possible. The peristaltic movement of the esophagus is an involuntary reflex; it takes place in response to the act of swallowing.



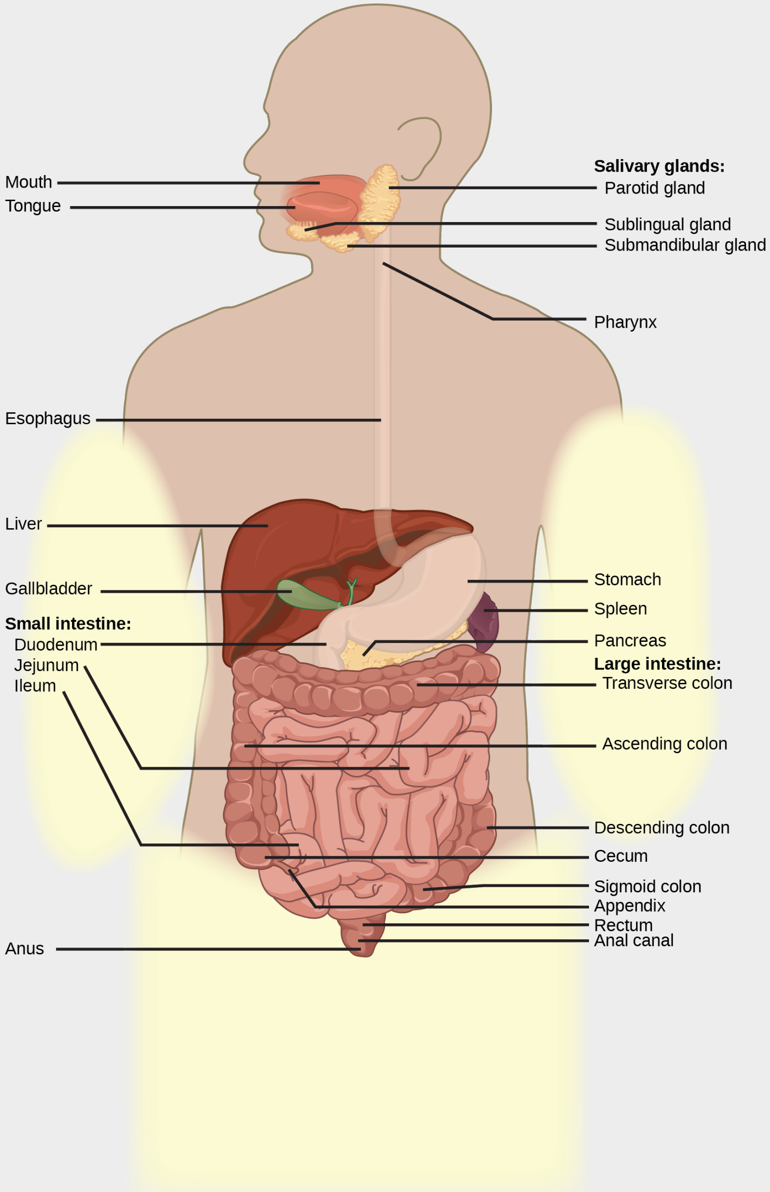
A ring-like muscle called a **sphincter** forms valves in the digestive system. The gastro-esophageal sphincter is located at the stomach end of the esophagus. In response to swallowing and the pressure exerted by the bolus of food, this sphincter opens, and the bolus enters the stomach. When there is no swallowing action, this sphincter is shut and prevents the contents of the stomach from traveling up the esophagus. Many animals have a true sphincter; however, in humans, there is no true sphincter, but the esophagus remains closed when there is no swallowing action. Acid reflux or “heartburn” occurs when the acidic digestive juices escape into the esophagus.

Stomach

A large part of digestion occurs in the stomach, shown in [link]. The **stomach** is a saclike organ that secretes gastric digestive juices. The pH in the stomach is between 1.5 and 2.5. This highly acidic environment is required for the chemical breakdown of food and the extraction of nutrients. When empty, the stomach is a rather small organ; however, it can expand to up to 20 times its resting size when filled with food. This characteristic is particularly useful for animals that need to eat when food is available.

Visual Connection

The human stomach has an extremely acidic environment where most of the protein gets digested. (credit: modification of work by Mariana Ruiz Villareal)



Which of the following statements about the digestive system is false?

1. Chyme is a mixture of food and digestive

juices that is produced in the stomach.
2. Food enters the large intestine before the small intestine.
3. In the small intestine, chyme mixes with bile, which emulsifies fats.
4. The stomach is separated from the small intestine by the pyloric sphincter.

The stomach is also the major site for protein digestion in animals other than ruminants. Protein digestion is mediated by an enzyme called pepsin in the stomach chamber. **Pepsin** is secreted by the chief cells in the stomach in an inactive form called **pepsinogen**. Pepsin breaks peptide bonds and cleaves proteins into smaller polypeptides; it also helps activate more pepsinogen, starting a positive feedback mechanism that generates more pepsin. Another cell type—parietal cells—secrete hydrogen and chloride ions, which combine in the lumen to form hydrochloric acid, the primary acidic component of the stomach juices. Hydrochloric acid helps to convert the inactive pepsinogen to pepsin. The highly acidic environment also kills many microorganisms in the food and, combined with the action of the enzyme pepsin, results in the hydrolysis of protein in the food. Chemical digestion is facilitated by the churning action of the stomach. Contraction and relaxation of smooth muscles mixes the stomach contents about every 20 minutes. The

partially digested food and gastric juice mixture is called **chyme**. Chyme passes from the stomach to the small intestine. Further protein digestion takes place in the small intestine. Gastric emptying occurs within two to six hours after a meal. Only a small amount of chyme is released into the small intestine at a time. The movement of chyme from the stomach into the small intestine is regulated by the pyloric sphincter.

When digesting protein and some fats, the stomach lining must be protected from getting digested by pepsin. There are two points to consider when describing how the stomach lining is protected. First, as previously mentioned, the enzyme pepsin is synthesized in the inactive form. This protects the chief cells, because pepsinogen does not have the same enzyme functionality of pepsin. Second, the stomach has a thick mucus lining that protects the underlying tissue from the action of the digestive juices. When this mucus lining is ruptured, ulcers can form in the stomach. Ulcers are open wounds in or on an organ caused by bacteria (*Helicobacter pylori*) when the mucus lining is ruptured and fails to reform.

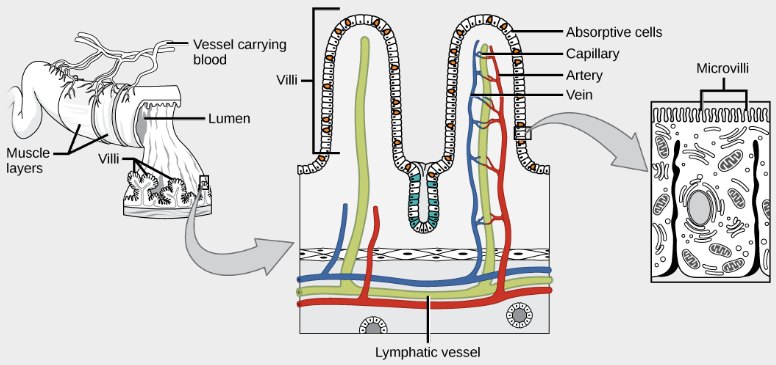
Small Intestine

Chyme moves from the stomach to the small intestine. The **small intestine** is the organ where the digestion of protein, fats, and carbohydrates is

completed. The small intestine is a long tube-like organ with a highly folded surface containing finger-like projections called the **villi**. The apical surface of each villus has many microscopic projections called microvilli. These structures, illustrated in [link], are lined with epithelial cells on the luminal side and allow for the nutrients to be absorbed from the digested food and absorbed into the bloodstream on the other side. The villi and microvilli, with their many folds, increase the surface area of the intestine and increase absorption efficiency of the nutrients. Absorbed nutrients in the blood are carried into the hepatic portal vein, which leads to the liver. There, the liver regulates the distribution of nutrients to the rest of the body and removes toxic substances, including drugs, alcohol, and some pathogens.

Visual Connection

Villi are folds on the small intestine lining that increase the surface area to facilitate the absorption of nutrients.



Which of the following statements about the small intestine is false?

1. Absorptive cells that line the small intestine have microvilli, small projections that increase surface area and aid in the absorption of food.
2. The inside of the small intestine has many folds, called villi.
3. Microvilli are lined with blood vessels as well as lymphatic vessels.
4. The inside of the small intestine is called the lumen.

The human small intestine is over 6m long and is divided into three parts: the duodenum, the jejunum, and the ileum. The "C-shaped," fixed part of the small intestine is called the **duodenum** and is shown in [link]. The duodenum is separated from the stomach by the pyloric sphincter which opens to allow chyme to move from the stomach to the duodenum. In the duodenum, chyme is mixed with

pancreatic juices in an alkaline solution rich in bicarbonate that neutralizes the acidity of chyme and acts as a buffer. Pancreatic juices also contain several digestive enzymes. Digestive juices from the pancreas, liver, and gallbladder, as well as from gland cells of the intestinal wall itself, enter the duodenum. **Bile** is produced in the liver and stored and concentrated in the gallbladder. Bile contains bile salts which emulsify lipids while the pancreas produces enzymes that catabolize starches, disaccharides, proteins, and fats. These digestive juices breakdown the food particles in the chyme into glucose, triglycerides, and amino acids. Some chemical digestion of food takes place in the duodenum. Absorption of fatty acids also takes place in the duodenum.

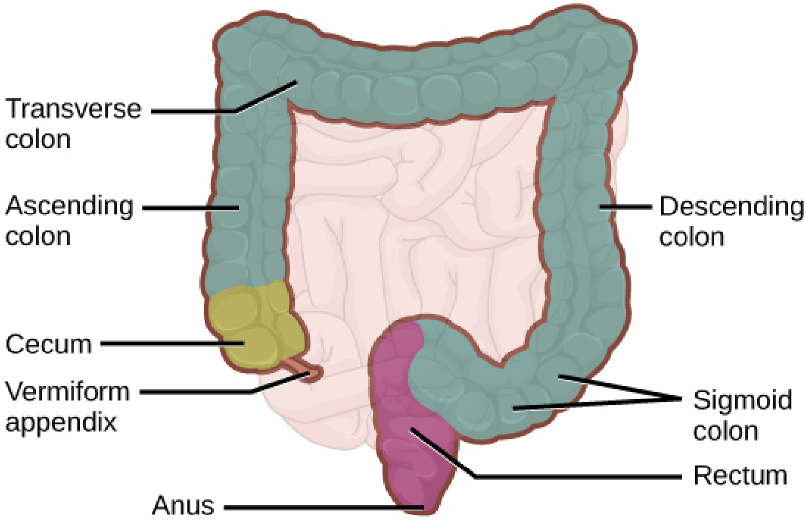
The second part of the small intestine is called the **jejunum**, shown in [link]. Here, hydrolysis of nutrients is continued while most of the carbohydrates and amino acids are absorbed through the intestinal lining. The bulk of chemical digestion and nutrient absorption occurs in the jejunum.

The **ileum**, also illustrated in [link] is the last part of the small intestine and here the bile salts and vitamins are absorbed into the bloodstream. The undigested food is sent to the colon from the ileum via peristaltic movements of the muscle. The ileum ends and the large intestine begins at the ileocecal

valve. The vermiform, “worm-like,” appendix is located at the ileocecal valve. The appendix of humans secretes no enzymes and has an insignificant role in immunity.

Large Intestine

The **large intestine**, illustrated in [link], reabsorbs the water from the undigested food material and processes the waste material. The human large intestine is much smaller in length compared to the small intestine but larger in diameter. It has three parts: the cecum, the colon, and the rectum. The cecum joins the ileum to the colon and is the receiving pouch for the waste matter. The colon is home to many bacteria or “intestinal flora” that aid in the digestive processes. The colon can be divided into four regions, the ascending colon, the transverse colon, the descending colon, and the sigmoid colon. The main functions of the colon are to extract the water and mineral salts from undigested food, and to store waste material. Carnivorous mammals have a shorter large intestine compared to herbivorous mammals due to their diet.



Rectum and Anus

The **rectum** is the terminal end of the large intestine, as shown in [link]. The primary role of the rectum is to store the feces until defecation. The feces are propelled using peristaltic movements during elimination. The **anus** is an opening at the far-end of the digestive tract and is the exit point for the waste material. Two sphincters between the rectum and anus control elimination: the inner sphincter is involuntary and the outer sphincter is voluntary.

Accessory Organs

The organs discussed above are the organs of the digestive tract through which food passes. Accessory organs are organs that add secretions (enzymes) that

catabolize food into nutrients. Accessory organs include salivary glands, the liver, the pancreas, and the gallbladder. The liver, pancreas, and gallbladder are regulated by hormones in response to the food consumed.

The **liver** is the largest internal organ in humans and it plays a very important role in digestion of fats and detoxifying blood. The liver produces bile, a digestive juice that is required for the breakdown of fatty components of the food in the duodenum. The liver also processes the vitamins and fats and synthesizes many plasma proteins.

The **pancreas** is another important gland that secretes digestive juices. The chyme produced from the stomach is highly acidic in nature; the pancreatic juices contain high levels of bicarbonate, an alkali that neutralizes the acidic chyme. Additionally, the pancreatic juices contain a large variety of enzymes that are required for the digestion of protein and carbohydrates.

The **gallbladder** is a small organ that aids the liver by storing bile and concentrating bile salts. When chyme containing fatty acids enters the duodenum, the bile is secreted from the gallbladder into the duodenum.

Section Summary

Different animals have evolved different types of digestive systems specialized to meet their dietary needs. Humans and many other animals have monogastric digestive systems with a single-chambered stomach. Birds have evolved a digestive system that includes a gizzard where the food is crushed into smaller pieces. This compensates for their inability to masticate. Ruminants that consume large amounts of plant material have a multi-chambered stomach that digests roughage. Pseudo-ruminants have similar digestive processes as ruminants but do not have the four-compartment stomach. Processing food involves ingestion (eating), digestion (mechanical and enzymatic breakdown of large molecules), absorption (cellular uptake of nutrients), and elimination (removal of undigested waste as feces).

Many organs work together to digest food and absorb nutrients. The mouth is the point of ingestion and the location where both mechanical and chemical breakdown of food begins. Saliva contains an enzyme called amylase that breaks down carbohydrates. The food bolus travels through the esophagus by peristaltic movements to the stomach. The stomach has an extremely acidic environment. An enzyme called pepsin digests protein in the stomach. Further digestion and absorption take place in the small intestine. The large intestine reabsorbs water from the undigested food and stores waste until elimination.

Visual Connection Questions

[link] Which of the following statements about the digestive system is false?

1. Chyme is a mixture of food and digestive juices that is produced in the stomach.
2. Food enters the large intestine before the small intestine.
3. In the small intestine, chyme mixes with bile, which emulsifies fats.
4. The stomach is separated from the small intestine by the pyloric sphincter.

[link] B

[link] Which of the following statements about the small intestine is false?

1. Absorptive cells that line the small intestine have microvilli, small projections that increase surface area and aid in the absorption of food.
2. The inside of the small intestine has many folds, called villi.
3. Microvilli are lined with blood vessels as

well as lymphatic vessels.
4. The inside of the small intestine is called the lumen.

[link] C

Review Questions

Which of the following is a pseudo-ruminant?

1. cow
2. pig
3. crow
4. horse

D

Which of the following statements is untrue?

1. Roughage takes a long time to digest.
2. Birds eat large quantities at one time so that they can fly long distances.
3. Cows do not have upper teeth.
4. In pseudo-ruminants, roughage is digested in the cecum.

B

The acidic nature of chyme is neutralized by ______.

1. potassium hydroxide
2. sodium hydroxide
3. bicarbonates
4. vinegar

C

The digestive juices from the liver are delivered to the ______.

1. stomach
2. liver
3. duodenum
4. colon

C

A scientist dissects a new species of animal. If the animal's digestive system has a single stomach with an extended small intestine, to which animal could the dissected specimen be closely related?

1. lion
2. snowshoe hare
3. earthworm
4. eagle

B

Critical Thinking Questions

How does the polygastric digestive system aid in digesting roughage?

Animals with a polygastric digestive system have a multi-chambered stomach. The four compartments of the stomach are called the rumen, reticulum, omasum, and abomasum. These chambers contain many microbes that breakdown the cellulose and ferment the ingested food. The abomasum is the “true” stomach and is the equivalent of a monogastric stomach chamber where gastric juices are secreted. The four-compartment gastric chamber provides larger space and the microbial support necessary for ruminants to digest plant material.

How do birds digest their food in the absence of teeth?

Birds have a stomach chamber called a gizzard. Here, the food is stored, soaked, and ground into finer particles, often using pebbles. Once this process is complete, the digestive juices take over in the proventriculus and continue the digestive process.

What is the role of the accessory organs in digestion?

Accessory organs play an important role in producing and delivering digestive juices to the intestine during digestion and absorption. Specifically, the salivary glands, liver, pancreas, and gallbladder play important roles. Malfunction of any of these organs can lead to disease states.

Explain how the villi and microvilli aid in absorption.

The villi and microvilli are folds on the surface of the small intestine. These folds increase the surface area of the intestine and provide more

area for the absorption of nutrients.

Name two components of the digestive system that perform mechanical digestion. Describe how mechanical digestion contributes to acquiring nutrients from food.

The stomach and the teeth both perform mechanical digestion, which is physically (as opposed to chemically) breaking the food into smaller components. This exposes a larger surface area for chemical digestion and release of nutrients. The teeth are vital to mastication, which breaks large bites of food down into smaller pieces that are easily swallowed. The stomach's muscle contractions churn the food to expose all particles to the acid and digestive enzymes.

Glossary

alimentary canal
: tubular digestive system with a mouth and anus

anus
: exit point for waste material

bile

digestive juice produced by the liver; important for digestion of lipids

bolus
: mass of food resulting from chewing action and wetting by saliva

carnivore
: animal that consumes animal flesh

chyme
: mixture of partially digested food and stomach juices

duodenum
: first part of the small intestine where a large part of digestion of carbohydrates and fats occurs

esophagus
: tubular organ that connects the mouth to the stomach

gallbladder
: organ that stores and concentrates bile

gastrovascular cavity
: digestive system consisting of a single opening

gizzard
: muscular organ that grinds food

herbivore
: animal that consumes a strictly plant diet

ileum
: last part of the small intestine; connects the small intestine to the large intestine; important for absorption of B-12

jejunum
: second part of the small intestine

large intestine
: digestive system organ that reabsorbs water from undigested material and processes waste matter

lipase
: enzyme that chemically breaks down lipids

liver
: organ that produces bile for digestion and processes vitamins and lipids

monogastric
: digestive system that consists of a single-chambered stomach

omnivore
: animal that consumes both plants and animals

pancreas
: gland that secretes digestive juices

pepsin

enzyme found in the stomach whose main role is protein digestion

pepsinogen

inactive form of pepsin

peristalsis

wave-like movements of muscle tissue

proventriculus

glandular part of a bird's stomach

rectum

area of the body where feces is stored until elimination

roughage

component of food that is low in energy and high in fiber

ruminant

animal with a stomach divided into four compartments

salivary amylase

enzyme found in saliva, which converts carbohydrates to maltose

small intestine

organ where digestion of protein, fats, and carbohydrates is completed

sphincter
: band of muscle that controls movement of materials throughout the digestive tract

stomach
: saclike organ containing acidic digestive juices

villi
: folds on the inner surface of the small intestine whose role is to increase absorption area

Nutrition and Energy Production

By the end of this section, you will be able to do the following:

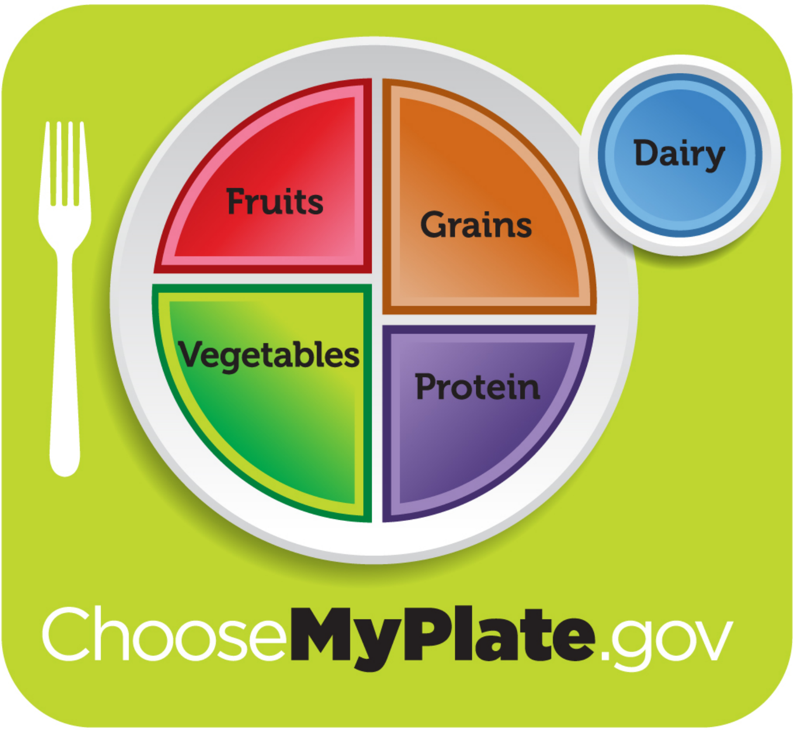
- Explain why an animal’s diet should be balanced and meet the needs of the body
- Define the primary components of food
- Describe the essential nutrients required for cellular function that cannot be synthesized by the animal body
- Explain how energy is produced through diet and digestion
- Describe how excess carbohydrates and energy are stored in the body

Given the diversity of animal life on our planet, it is not surprising that the animal diet would also vary substantially. The animal diet is the source of materials needed for building DNA and other complex molecules needed for growth, maintenance, and reproduction; collectively these processes are called biosynthesis. The diet is also the source of materials for ATP production in the cells. The diet must be balanced to provide the minerals and vitamins that are required for cellular function.

For humans, a balanced diet includes fruits, vegetables, grains, and protein. (credit: USDA) A healthy diet should include a variety of foods to ensure that needs for essential nutrients are met. (credit: Keith Weller, USDA ARS)

Food Requirements

What are the fundamental requirements of the animal diet? The animal diet should be well balanced and provide nutrients required for bodily function and the minerals and vitamins required for maintaining structure and regulation necessary for good health and reproductive capability. These requirements for a human are illustrated graphically in [link]



Link to Learning

The first step in ensuring that you are meeting the

food requirements of your body is an awareness of the food groups and the nutrients they provide. To learn more about each food group and the recommended daily amounts, explore this interactive site by the United States Department of Agriculture.

Everyday Connection

Let's Move! Campaign

Obesity is a growing epidemic and the rate of obesity among children is rapidly rising in the United States. To combat childhood obesity and ensure that children get a healthy start in life, first lady Michelle Obama has launched the Let's Move! campaign. The goal of this campaign is to educate parents and caregivers on providing healthy nutrition and encouraging active lifestyles to future generations. This program aims to involve the entire community, including parents, teachers, and healthcare providers to ensure that children have access to healthy foods—more fruits, vegetables, and whole grains—and consume fewer calories from processed foods. Another goal is to ensure that children get physical activity. With the increase in television viewing and stationary pursuits such as video games, sedentary lifestyles have become the norm. Learn more at https://letsmove.obamawhitehouse.archives.gov.

Organic Precursors

The organic molecules required for building cellular material and tissues must come from food. Carbohydrates or sugars are the primary source of organic carbons in the animal body. During digestion, digestible carbohydrates are ultimately broken down into glucose and used to provide energy through metabolic pathways. Complex carbohydrates, including polysaccharides, can be broken down into glucose through biochemical modification; however, humans do not produce the enzyme cellulase and lack the ability to derive glucose from the polysaccharide cellulose. In humans, these molecules provide the fiber required for moving waste through the large intestine and a healthy colon. The intestinal flora in the human gut are able to extract some nutrition from these plant fibers. The excess sugars in the body are converted into glycogen and stored in the liver and muscles for later use. Glycogen stores are used to fuel prolonged exertions, such as long-distance running, and to provide energy during food shortage. Excess glycogen can be converted to fats, which are stored in the lower layer of the skin of mammals for insulation and energy storage. Excess digestible carbohydrates are stored by mammals in order to survive famine and aid in mobility.

Another important requirement is that of nitrogen. Protein catabolism provides a source of organic

nitrogen. Amino acids are the building blocks of proteins and protein breakdown provides amino acids that are used for cellular function. The carbon and nitrogen derived from these become the building block for nucleotides, nucleic acids, proteins, cells, and tissues. Excess nitrogen must be excreted as it is toxic. Fats add flavor to food and promote a sense of satiety or fullness. Fatty foods are also significant sources of energy because one gram of fat contains nine calories. Fats are required in the diet to aid the absorption of fat-soluble vitamins and the production of fat-soluble hormones.

Essential Nutrients

While the animal body can synthesize many of the molecules required for function from the organic precursors, there are some nutrients that need to be consumed from food. These nutrients are termed **essential nutrients**, meaning they must be eaten, and the body cannot produce them.

The omega-3 alpha-linolenic acid and the omega-6 linoleic acid are essential fatty acids needed to make some membrane phospholipids. **Vitamins** are another class of essential organic molecules that are required in small quantities for many enzymes to function and, for this reason, are considered to be coenzymes. Absence or low levels of vitamins can have a dramatic effect on health, as outlined in

[link] and [link]. Both fat-soluble and water-soluble vitamins must be obtained from food. **Minerals**, listed in [link], are inorganic essential nutrients that must be obtained from food. Among their many functions, minerals help in structure and regulation and are considered cofactors. Certain amino acids also must be procured from food and cannot be synthesized by the body. These amino acids are the “essential” amino acids. The human body can synthesize only 11 of the 20 required amino acids; the rest must be obtained from food. The essential amino acids are listed in [link].

Water-soluble Essential Vitamins

Vitamin	Function	Deficiencies Can Lead To	Sources
Vitamin B1 (Thiamine)	Needed by the body to process lipids, proteins, and carbohydrates; coenzyme removes CO_2	Muscle weakness, Beriberi: reduced heart function, CNS problems	Milk, meat, dried beans, whole grains

	from organic compounds		
Vitamin B2 (Riboflavin)	Takes an active role in metabolism, aiding in the conversion of food to energy (FAD and FMN)	Cracks or sores on the outer surface of the lips (cheliosis); inflammation and redness of the tongue; moist, scaly skin inflammation (seborrheic dermatitis)	Meat, eggs, enriched grains, vegetables
Vitamin B3 (Niacin)	Used by the body to release energy from carbohydrates and to process alcohol; required for the synthesis of sex hormones; component of coenzyme NAD+ and	Pellagra, which can result in dermatitis, diarrhea, dementia, and death	Meat, eggs, grains, nuts, potatoes

	NADP+		
Vitamin B5 (Pantothenic acid)	Assists in producing energy from foods (lipids, in particular); component of coenzyme A	Fatigue, poor coordination, retarded growth, numbness, tingling of hands and feet	Meat, whole grains, milk, fruits, vegetables
Vitamin B6 (Pyridoxine)	The principal vitamin for processing amino acids and lipids; also helps convert nutrients into energy	Irritability, depression, confusion, mouth sores or ulcers, anemia, muscular twitching	Meat, dairy products, whole grains, orange juice
Vitamin B7 (Biotin)	Used in energy and amino acid metabolism, fat synthesis, and fat breakdown; helps the body use blood sugar	Hair loss, dermatitis, depression, numbness and tingling in the extremities; neuromuscular disorders	Meat, eggs, legumes and other vegetables
Vitamin B9	Assists the	Deficiency	Leafy green

(Folic acid)	normal development of cells, especially during fetal development; helps metabolize nucleic and amino acids	during pregnancy is associated with birth defects, such as neural tube defects and anemia	vegetables, whole wheat, fruits, nuts, legumes
Vitamin B12 (Cobalamin)	Maintains healthy nervous system and assists with blood cell formation; coenzyme in nucleic acid metabolism	Anemia, neurological disorders, numbness, loss of balance	Meat, eggs, animal products
Vitamin C (Ascorbic acid)	Helps maintain connective tissue: bone, cartilage, and dentin; boosts the immune system	Scurvy, which results in bleeding, hair and tooth loss; joint pain and swelling; delayed wound	Citrus fruits, broccoli, tomatoes, red sweet bell peppers

healing

Fat-soluble Essential Vitamins

Vitamin	**Function**	**Deficiencies Can Lead To**	**Sources**
Vitamin A (Retinol)	Critical to the development of bones, teeth, and skin; helps maintain eyesight, enhances the immune system, fetal development, gene expression	Night-blindness, skin disorders, impaired immunity	Dark green leafy vegetables, yellow-orange vegetables, fruits, milk, butter
Vitamin D	Critical for calcium absorption for bone development and strength;	Rickets, osteomalacia, immunity	Cod liver oil, milk, egg yolk

	maintains a stable nervous system; maintains a normal and strong heartbeat; helps in blood clotting		
Vitamin E (Tocopherol)	Lessens oxidative damage of cells and prevents lung damage from pollutants; vital to the immune system	Deficiency is rare; anemia, nervous system degeneration	Wheat germ oil, unrefined vegetable oils, nuts, seeds, grains
Vitamin K (Phylloquinone)	Essential to blood clotting	Bleeding and easy bruising	Leafy green vegetables, tea

Minerals and Their Function in the Human Body

Mineral	Function	Deficiencies Can Lead To	Sources
*Calcium	Needed for muscle and neuron function; heart health; builds bone and supports	Osteoporosis, rickets, muscle spasms, impaired growth	Milk, yogurt, fish, green leafy vegetables, legumes

	synthesis and function of blood cells; nerve function		
*Chlorine	Needed for production of hydrochloric acid (HCl) in the stomach and nerve function; osmotic balance	Muscle cramps, mood disturbances, reduced appetite	Table salt
Copper (trace amounts)	Required component of many redox enzymes, including cytochrome c oxidase; cofactor for hemoglobin synthesis	Copper deficiency is rare	Liver, oysters, cocoa, chocolate, sesame, nuts
Iodine	Required for the synthesis of thyroid hormones	Goiter	Seafood, iodized salt, dairy products
Iron	Required for	Anemia,	Red meat,

	many proteins and enzymes, notably hemoglobin, to prevent anemia	which causes poor concentration, fatigue, and poor immune function	leafy green vegetables, fish (tuna, salmon), eggs, dried fruits, beans, whole grains
*Magnesium	Required cofactor for ATP formation; bone formation; normal membrane functions; muscle function	Mood disturbances, muscle spasms	Whole grains, leafy green vegetables
Manganese (trace amounts)	A cofactor in enzyme functions; trace amounts are required	Manganese deficiency is rare	Common in most foods
Molybdenum (trace amounts)	Acts as a cofactor for three essential enzymes in humans: sulfite	Molybdenum deficiency is rare	

	oxidase, xanthine oxidase, and aldehyde oxidase		
*Phosphorus	A component of bones and teeth; helps regulate acid-base balance; nucleotide synthesis	Weakness, bone abnormalities, calcium loss	Milk, hard cheese, whole grains, meats
*Potassium	Vital for muscles, heart, and nerve function	Cardiac rhythm disturbance, muscle weakness	Legumes, potato skin, tomatoes, bananas
Selenium (trace amounts)	A cofactor essential to activity of antioxidant enzymes like glutathione peroxidase; trace amounts are required	Selenium deficiency is rare	Common in most foods
*Sodium	Systemic electrolyte	Muscle cramps,	Table salt

	required for many functions; acid-base balance; water balance; nerve function	fatigue, reduced appetite	
Zinc (trace amounts)	Required for several enzymes such as carboxypeptidase, liver alcohol dehydrogenase, and carbonic anhydrase	Anemia, poor wound healing, can lead to short stature	Common in most foods
*Greater than 200mg/ day required			

Essential Amino Acids	
Amino acids that must be consumed	Amino acids anabolized by the body
isoleucine	alanine
leucine	selenocysteine

lysine	aspartate
methionine	cysteine
phenylalanine	glutamate
tryptophan	glycine
valine	proline
histidine*	serine
threonine	tyrosine
arginine*	asparagine
*The human body can synthesize histidine and arginine, but not in the quantities required, especially for growing children.	

Food Energy and ATP

Animals need food to obtain energy and maintain homeostasis. Homeostasis is the ability of a system to maintain a stable internal environment even in the face of external changes to the environment. For example, the normal body temperature of humans is 37°C (98.6°F). Humans maintain this temperature even when the external temperature is hot or cold. It takes energy to maintain this body temperature, and animals obtain this energy from food.

The primary source of energy for animals is carbohydrates, mainly glucose. Glucose is called the

body's fuel. The digestible carbohydrates in an animal's diet are converted to glucose molecules through a series of catabolic chemical reactions.

Adenosine triphosphate, or ATP, is the primary energy currency in cells; ATP stores energy in phosphate ester bonds. ATP releases energy when the phosphodiester bonds are broken and ATP is converted to ADP and a phosphate group. ATP is produced by the oxidative reactions in the cytoplasm and mitochondrion of the cell, where carbohydrates, proteins, and fats undergo a series of metabolic reactions collectively called cellular respiration. For example, glycolysis is a series of reactions in which glucose is converted to pyruvic acid and some of its chemical potential energy is transferred to NADH and ATP.

ATP is required for all cellular functions. It is used to build the organic molecules that are required for cells and tissues; it provides energy for muscle contraction and for the transmission of electrical signals in the nervous system. When the amount of ATP is available in excess of the body's requirements, the liver uses the excess ATP and excess glucose to produce molecules called glycogen. Glycogen is a polymeric form of glucose and is stored in the liver and skeletal muscle cells. When blood sugar drops, the liver releases glucose from stores of glycogen. Skeletal muscle converts glycogen to glucose during intense exercise. The

process of converting glucose and excess ATP to glycogen and the storage of excess energy is an evolutionarily important step in helping animals deal with mobility, food shortages, and famine.

Everyday Connection

Obesity

Obesity is a major health concern in the United States, and there is a growing focus on reducing obesity and the diseases it may lead to, such as type-2 diabetes, cancers of the colon and breast, and cardiovascular disease. How does the food consumed contribute to obesity?

Fatty foods are calorie-dense, meaning that they have more calories per unit mass than carbohydrates or proteins. One gram of carbohydrates has four calories, one gram of protein has four calories, and one gram of fat has nine calories. Animals tend to seek lipid-rich food for their higher energy content.

The signals of hunger ("time to eat") and satiety ("time to stop eating") are controlled in the hypothalamus region of the brain. Foods that are rich in fatty acids tend to promote satiety more than foods that are rich only in carbohydrates.

Excess carbohydrate and ATP are used by the liver to synthesize glycogen. The pyruvate produced during glycolysis is used to synthesize fatty acids. When there is more glucose in the body than

required, the resulting excess pyruvate is converted into molecules that eventually result in the synthesis of fatty acids within the body. These fatty acids are stored in adipose cells—the fat cells in the mammalian body whose primary role is to store fat for later use.

It is important to note that some animals benefit from obesity. Polar bears and seals need body fat for insulation and to keep them from losing body heat during Arctic winters. When food is scarce, stored body fat provides energy for maintaining homeostasis. Fats prevent famine in mammals, allowing them to access energy when food is not available on a daily basis; fats are stored when a large kill is made or lots of food is available.

Section Summary

Animal diet should be balanced and meet the needs of the body. Carbohydrates, proteins, and fats are the primary components of food. Some essential nutrients are required for cellular function but cannot be produced by the animal body. These include vitamins, minerals, some fatty acids, and some amino acids. Food intake in more than necessary amounts is stored as glycogen in the liver and muscle cells, and in fat cells. Excess adipose

storage can lead to obesity and serious health problems. ATP is the energy currency of the cell and is obtained from the metabolic pathways. Excess carbohydrates and energy are stored as glycogen in the body.

Review Questions

Which of the following statements is not true?

1. Essential nutrients can be synthesized by the body.
2. Vitamins are required in small quantities for bodily function.
3. Some amino acids can be synthesized by the body, while others need to be obtained from diet.
4. Vitamins come in two categories: fat-soluble and water-soluble.

A

Which of the following is a water-soluble vitamin?

1. vitamin A
2. vitamin E

3. vitamin K
4. vitamin C

D

What is the primary fuel for the body?

1. carbohydrates
2. lipids
3. protein
4. glycogen

A

Excess glucose is stored as ______.

1. fat
2. glucagon
3. glycogen
4. it is not stored in the body

C

Many distance runners “carb load” the day before a big race. How does this eating strategy provide an advantage to the runner?

1. The carbohydrates cause the release of insulin.
2. The excess carbohydrates are converted to fats, which have a higher calorie density.
3. The glucose from the carbohydrates lets the muscles make excess ATP overnight.
4. The excess carbohydrates can be stored in the muscles as glycogen.

D

Critical Thinking Questions

What are essential nutrients?

Essential nutrients are those nutrients that must be obtained from the diet because they cannot be produced by the body. Vitamins and minerals are examples of essential nutrients.

What is the role of minerals in maintaining good health?

Minerals—such as potassium, sodium, and calcium—are required for the functioning of

many cellular processes, including muscle contraction and nerve conduction. While minerals are required in trace amounts, not having minerals in the diet can be potentially harmful.

Discuss why obesity is a growing epidemic.

In the United States, obesity, particularly childhood obesity, is a growing concern. Some of the contributors to this situation include sedentary lifestyles and consuming more processed foods and less fruits and vegetables. As a result, even young children who are obese can face health concerns.

There are several nations where malnourishment is a common occurrence. What may be some of the health challenges posed by malnutrition?

Malnutrition, often in the form of not getting enough calories or not enough of the essential nutrients, can have severe consequences. Many malnourished children have vision and dental problems, and over the years may develop many serious health problems.

Generally describe how a piece of bread can power your legs as you walk up a flight of stairs.

A piece of bread is eaten and converted into chemical energy. The bread is broken down in the mouth by mastication and salivary enzymes, then transferred to the stomach for further digestion. After digestion by the acid and digestive enzymes in the stomach, the macromolecules that made up the bread move into the small intestine. In the small intestine, the carbohydrates from the bread are absorbed through the microvilli into the bloodstream. In muscle cells in the legs, the carbohydrates can be broken down into glucose, and then used for cellular respiration to create ATP. The muscles in the leg then use the ATP to perform the mechanical work needed to climb a flight of stairs.

In the 1990s fat-free foods became popular among people trying to lose weight. However, many dieticians now conclude that the fat-free trend made people less healthy and heavier. Describe how this could occur.

Fats are an essential component of a healthy diet, and needed by the body to function. Fats

are essential for many processes, including the absorption of fat-soluble vitamins and production of some hormones. Fats also send a satiation signal to the brain that regulates hunger. Without fats in their diets many people may have actually consumed more calories, which would have resulted in weight gain.

Glossary

essential nutrient
: nutrient that cannot be synthesized by the body; it must be obtained from food

mineral
: inorganic, elemental molecule that carries out important roles in the body

vitamin
: organic substance necessary in small amounts to sustain life

Digestive System Processes

By the end of this section, you will be able to do the following:

- Describe the process of digestion
- Detail the steps involved in digestion and absorption
- Define elimination
- Explain the role of both the small and large intestines in absorption

Obtaining nutrition and energy from food is a multistep process. For true animals, the first step is ingestion, the act of taking in food. This is followed by digestion, absorption, and elimination. In the following sections, each of these steps will be discussed in detail.

Ingestion

The large molecules found in intact food cannot pass through the cell membranes. Food needs to be broken into smaller particles so that animals can harness the nutrients and organic molecules. The first step in this process is **ingestion**. Ingestion is the process of taking in food through the mouth. In vertebrates, the teeth, saliva, and tongue play important roles in mastication (preparing the food into bolus). While the food is being mechanically

broken down, the enzymes in saliva begin to chemically process the food as well. The combined action of these processes modifies the food from large particles to a soft mass that can be swallowed and can travel the length of the esophagus.

Digestion of carbohydrates is performed by several enzymes. Starch and glycogen are broken down into glucose by amylase and maltase. Sucrose (table sugar) and lactose (milk sugar) are broken down by sucrase and lactase, respectively. Protein digestion is a multistep process that begins in the stomach and continues through the intestines. Lipids are digested and absorbed in the small intestine.

Digestion and Absorption

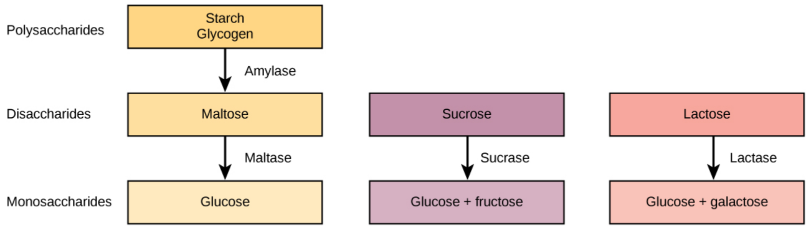
Digestion is the mechanical and chemical breakdown of food into small organic fragments. It is important to breakdown macromolecules into smaller fragments that are of suitable size for absorption across the digestive epithelium. Large, complex molecules of proteins, polysaccharides, and lipids must be reduced to simpler particles such as simple sugar before they can be absorbed by the digestive epithelial cells. Different organs play specific roles in the digestive process. The animal diet needs carbohydrates, protein, and fat, as well as vitamins and inorganic components for nutritional balance. How each of these components is digested is discussed in the following sections.

Carbohydrates

The digestion of carbohydrates begins in the mouth. The salivary enzyme amylase begins the breakdown of food starches into maltose, a disaccharide. As the bolus of food travels through the esophagus to the stomach, no significant digestion of carbohydrates takes place. The esophagus produces no digestive enzymes but does produce mucous for lubrication. The acidic environment in the stomach stops the action of the amylase enzyme.

The next step of carbohydrate digestion takes place in the duodenum. Recall that the chyme from the stomach enters the duodenum and mixes with the digestive secretion from the pancreas, liver, and gallbladder. Pancreatic juices also contain amylase, which continues the breakdown of starch and glycogen into maltose, a disaccharide. The disaccharides are broken down into monosaccharides by enzymes called **maltases**, **sucrases**, and **lactases**, which are also present in the brush border of the small intestinal wall. Maltase breaks down maltose into glucose. Other disaccharides, such as sucrose and lactose are broken down by sucrase and lactase, respectively. Sucrase breaks down sucrose (or “table sugar”) into glucose and fructose, and lactase breaks down lactose (or “milk sugar”) into glucose and galactose. The monosaccharides (glucose) thus produced are absorbed and then can be used in metabolic

pathways to harness energy. The monosaccharides are transported across the intestinal epithelium into the bloodstream to be transported to the different cells in the body. The steps in carbohydrate digestion are summarized in [link] and [link].



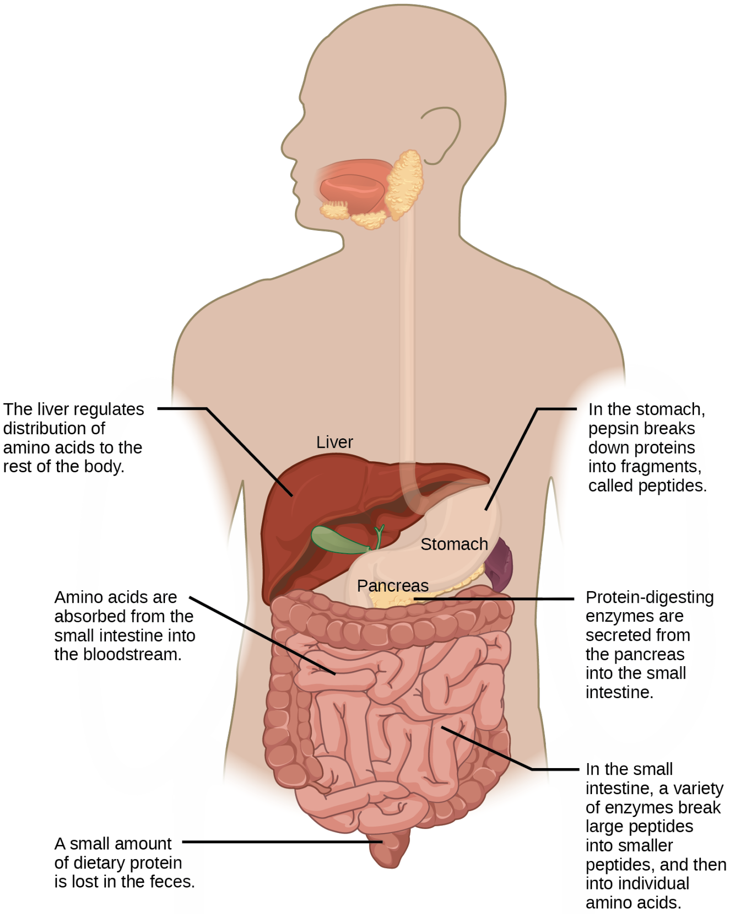
Digestion of Carbohydrates

Enzyme	Produced By	Site of Action	Substrate Acting On	End Products
Salivary amylase	Salivary glands	Mouth	Polysaccharides (Starch)	Disaccharides (maltose), oligosaccharides
Pancreatic amylase	Pancreas	Small intestine	Polysaccharides (starch)	Disaccharides (maltose), monosaccharides
Oligosaccharidases	Lining of the intestine; brush	Small intestine	Disaccharides	Monosaccharides (e.g., glucose, fructose,

	border membrane		galactose)

Protein

A large part of protein digestion takes place in the stomach. The enzyme pepsin plays an important role in the digestion of proteins by breaking down the intact protein to peptides, which are short chains of four to nine amino acids. In the duodenum, other enzymes—**trypsin**, **elastase**, and **chymotrypsin**—act on the peptides reducing them to smaller peptides. Trypsin elastase, carboxypeptidase, and chymotrypsin are produced by the pancreas and released into the duodenum where they act on the chyme. Further breakdown of peptides to single amino acids is aided by enzymes called peptidases (those that breakdown peptides). Specifically, **carboxypeptidase**, **dipeptidase**, and **aminopeptidase** play important roles in reducing the peptides to free amino acids. The amino acids are absorbed into the bloodstream through the small intestines. The steps in protein digestion are summarized in [link] and [link].



The liver regulates distribution of amino acids to the rest of the body.
Liver
In the stomach, pepsin breaks down proteins into fragments, called peptides.
Stomach
Pancreas
Amino acids are absorbed from the small intestine into the bloodstream.
Protein-digesting enzymes are secreted from the pancreas into the small intestine.
In the small intestine, a variety of enzymes break large peptides into smaller peptides, and then into individual amino acids.
A small amount of dietary protein is lost in the feces.

Digestion

of Protein

Enzyme	Produced By	Site of Action	Substrate Acting On	End Products
Pepsin	Stomach chief cells	Stomach	Proteins	Peptides
Trypsin Elastase Chymotrypsin	Pancreas	Small intestine	Proteins	Peptides
Carboxypeptidase	Pancreas	Small intestine	Peptides	Amino acids and peptides
Aminopeptidase Dipeptidase	Lining of intestine	Small intestine	Peptides	Amino acids

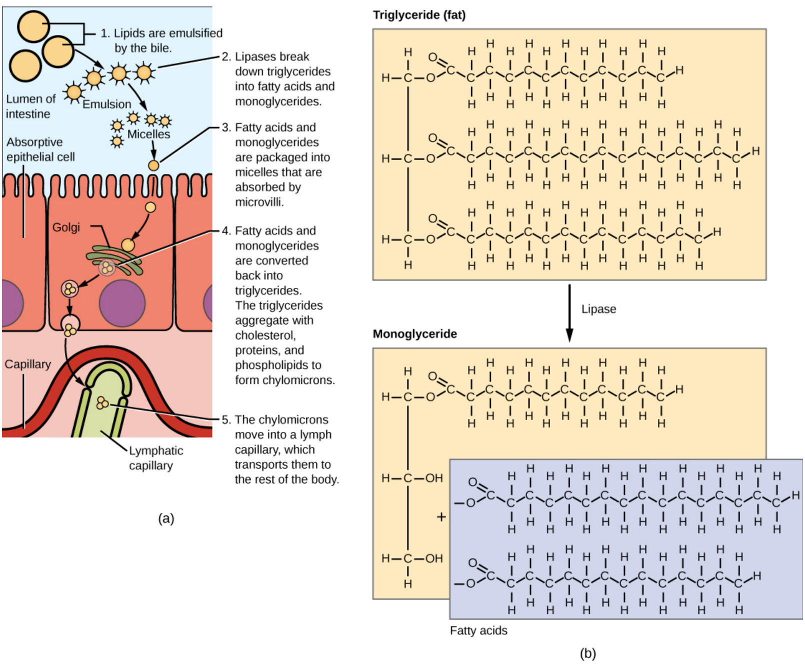
Lipids

Lipid digestion begins in the stomach with the aid of lingual lipase and gastric lipase. However, the bulk of lipid digestion occurs in the small intestine due to pancreatic lipase. When chyme enters the duodenum, the hormonal responses trigger the release of bile, which is produced in the liver and stored in the gallbladder. Bile aids in the digestion of lipids, primarily triglycerides by emulsification. Emulsification is a process in which large lipid globules are broken down into several small lipid globules. These small globules are more widely distributed in the chyme rather than forming large aggregates. Lipids are hydrophobic substances: in the presence of water, they will aggregate to form

globules to minimize exposure to water. Bile contains bile salts, which are amphipathic, meaning they contain hydrophobic and hydrophilic parts. Thus, the bile salts hydrophilic side can interface with water on one side and the hydrophobic side interfaces with lipids on the other. By doing so, bile salts emulsify large lipid globules into small lipid globules.

Why is emulsification important for digestion of lipids? Pancreatic juices contain enzymes called lipases (enzymes that breakdown lipids). If the lipid in the chyme aggregates into large globules, very little surface area of the lipids is available for the lipases to act on, leaving lipid digestion incomplete. By forming an emulsion, bile salts increase the available surface area of the lipids many fold. The pancreatic lipases can then act on the lipids more efficiently and digest them, as detailed in [link]. Lipases breakdown the lipids into fatty acids and glycerides. These molecules can pass through the plasma membrane of the cell and enter the epithelial cells of the intestinal lining. The bile salts surround long-chain fatty acids and monoglycerides forming tiny spheres called micelles. The micelles move into the brush border of the small intestine absorptive cells where the long-chain fatty acids and monoglycerides diffuse out of the micelles into the absorptive cells leaving the micelles behind in the chyme. The long-chain fatty acids and monoglycerides recombine in the absorptive cells to

form triglycerides, which aggregate into globules and become coated with proteins. These large spheres are called **chylomicrons**. Chylomicrons contain triglycerides, cholesterol, and other lipids and have proteins on their surface. The surface is also composed of the hydrophilic phosphate "heads" of phospholipids. Together, they enable the chylomicron to move in an aqueous environment without exposing the lipids to water. Chylomicrons leave the absorptive cells via exocytosis. Chylomicrons enter the lymphatic vessels, and then enter the blood in the subclavian vein.



Vitamins

Vitamins can be either water-soluble or lipid-

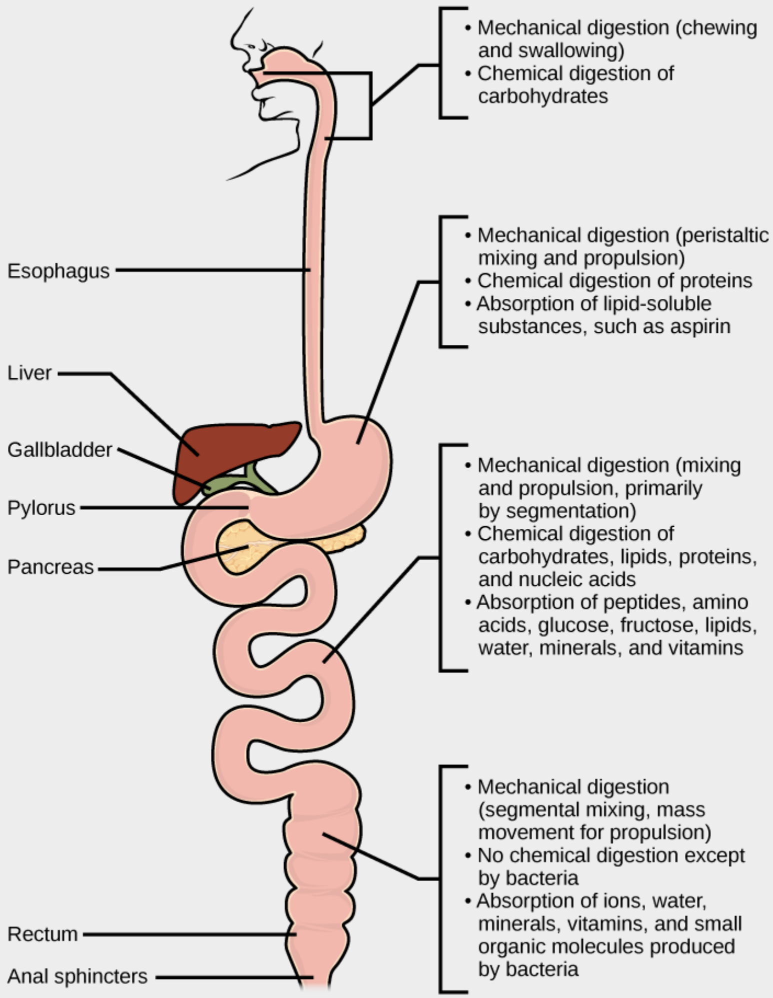
soluble. Fat soluble vitamins are absorbed in the same manner as lipids. It is important to consume some amount of dietary lipid to aid the absorption of lipid-soluble vitamins. Water-soluble vitamins can be directly absorbed into the bloodstream from the intestine.

Link to Learning

This website has an overview of the digestion of protein, fat, and carbohydrates.

Visual Connection

Mechanical and chemical digestion of food takes place in many steps, beginning in the mouth and ending in the rectum.



Which of the following statements about digestive processes is true?

1. Amylase, maltase, and lactase in the mouth digest carbohydrates.
2. Trypsin and lipase in the stomach digest protein.
3. Bile emulsifies lipids in the small intestine.
4. No food is absorbed until the small intestine.

Elimination

The final step in digestion is the elimination of undigested food content and waste products. The undigested food material enters the colon, where most of the water is reabsorbed. Recall that the colon is also home to the microflora called “intestinal flora” that aid in the digestion process. The semi-solid waste is moved through the colon by peristaltic movements of the muscle and is stored in the rectum. As the rectum expands in response to storage of fecal matter, it triggers the neural signals required to set up the urge to eliminate. The solid waste is eliminated through the anus using peristaltic movements of the rectum.

Common Problems with Elimination

Diarrhea and constipation are some of the most common health concerns that affect digestion. Constipation is a condition where the feces are hardened because of excess water removal in the colon. In contrast, if enough water is not removed from the feces, it results in diarrhea. Many bacteria, including the ones that cause cholera, affect the proteins involved in water reabsorption in the colon and result in excessive diarrhea.

Emesis

Emesis, or vomiting, is elimination of food by forceful expulsion through the mouth. It is often in response to an irritant that affects the digestive tract, including but not limited to viruses, bacteria, emotions, sights, and food poisoning. This forceful expulsion of the food is due to the strong contractions produced by the stomach muscles. The process of emesis is regulated by the medulla.

Section Summary

Digestion begins with ingestion, where the food is taken in the mouth. Digestion and absorption take place in a series of steps with special enzymes playing important roles in digesting carbohydrates, proteins, and lipids. Elimination describes removal of undigested food contents and waste products from the body. While most absorption occurs in the small intestines, the large intestine is responsible for the final removal of water that remains after the absorptive process of the small intestines. The cells that line the large intestine absorb some vitamins as well as any leftover salts and water. The large intestine (colon) is also where feces is formed.

Visual Connection Questions

[link] Which of the following statements about digestive processes is true?

1. Amylase, maltase, and lactase in the mouth digest carbohydrates.
2. Trypsin and lipase in the stomach digest protein.
3. Bile emulsifies lipids in the small intestine.
4. No food is absorbed until the small intestine.

[link] C

Review Questions

Where does the majority of protein digestion take place?

1. stomach
2. duodenum
3. mouth
4. jejunum

A

Lipases are enzymes that breakdown ______.

1. disaccharides
2. lipids
3. proteins
4. cellulose

B

Which of the following conditions is most likely to cause constipation?

1. bacterial infection
2. dehydration
3. ulcer
4. excessive cellulose consumption

B

Critical Thinking Questions

Explain why some dietary lipid is a necessary part of a balanced diet.

Lipids add flavor to food and promote a sense

of satiety or fullness. Fatty foods are sources of high energy; one gram of lipid contains nine calories. Lipids are also required in the diet to aid the absorption of lipid-soluble vitamins and for the production of lipid-soluble hormones.

The gut microbiome (the bacterial colonies in the intestines) have become a popular area of study in biomedical research. How could varying gut microbiomes impact a person's nutrition?

The gut microbiome includes all the bacteria that aid in chemical digestion in the intestines. Changing its composition can change the way that food is digested since not all bacteria have the same macromolecule-digesting enzymes. Additionally, changes in gut microbiome can lead to the establishment of pathogenic bacteria populations that cause inflammation in the gut or other disease.

Many mammals become ill if they drink milk as adults even though they could consume it as babies. What causes this digestive issue?

As mammals wean from their mothers they stop drinking milk. Since they stop consuming the

sugar lactose their bodies conserve resources by no longer making the enzyme lactase. If the animals then consume lactose at some point in the future their digestive system cannot break the lactose molecules into glucose and galactose for absorption. When gut bacteria further along the digestive tract interact with the lactose molecules it causes symptoms of lactose intolerance.

Glossary

aminopeptidase
: protease that breaks down peptides to single amino acids; secreted by the brush border of small intestine

carboxypeptidase
: protease that breaks down peptides to single amino acids; secreted by the brush border of the small intestine

chylomicron
: small lipid globule

chymotrypsin
: pancreatic protease

digestion
: mechanical and chemical breakdown of food into small organic fragments

dipeptidase
: protease that breaks down peptides to single amino acids; secreted by the brush border of small intestine

elastase
: pancreatic protease

ingestion
: act of taking in food

lactase
: enzyme that breaks down lactose into glucose and galactose

maltase
: enzyme that breaks down maltose into glucose

sucrase
: enzyme that breaks down sucrose into glucose and fructose

trypsin
: pancreatic protease that breaks down protein

Digestive System Regulation
By the end of this section, you will be able to do the following:

- Discuss the role of neural regulation in digestive processes
- Explain how hormones regulate digestion

The brain is the control center for the sensation of hunger and satiety. The functions of the digestive system are regulated through neural and hormonal responses.
Seeing a plate of food triggers the secretion of saliva in the mouth and the production of HCL in the stomach. (credit: Kelly Bailey)

Neural Responses to Food

In reaction to the smell, sight, or thought of food, like that shown in [link], the first response is that of salivation. The salivary glands secrete more saliva in response to stimulation by the autonomic nervous system triggered by food in preparation for digestion. Simultaneously, the stomach begins to produce hydrochloric acid to digest the food. Recall that the peristaltic movements of the esophagus and other organs of the digestive tract are under the control of the brain. The brain prepares these muscles for movement as well. When the stomach is full, the part of the brain that detects satiety signals

fullness. There are three overlapping phases of gastric control—the cephalic phase, the gastric phase, and the intestinal phase—each requires many enzymes and is under neural control as well.

Digestive Phases

The response to food begins even before food enters the mouth. The first phase of ingestion, called the **cephalic phase**, is controlled by the neural response to the stimulus provided by food. All aspects—such as sight, sense, and smell—trigger the neural responses resulting in salivation and secretion of gastric juices. The gastric and salivary secretion in the cephalic phase can also take place due to the thought of food. Right now, if you think about a piece of chocolate or a crispy potato chip, the

increase in salivation is a cephalic phase response to the thought. The central nervous system prepares the stomach to receive food.

The **gastric phase** begins once the food arrives in the stomach. It builds on the stimulation provided during the cephalic phase. Gastric acids and enzymes process the ingested materials. The gastric phase is stimulated by (1) distension of the stomach, (2) a decrease in the pH of the gastric contents, and (3) the presence of undigested material. This phase consists of local, hormonal, and neural responses. These responses stimulate secretions and powerful contractions.

The **intestinal phase** begins when chyme enters the small intestine triggering digestive secretions. This phase controls the rate of gastric emptying. In addition to gastrin emptying, when chyme enters the small intestine, it triggers other hormonal and neural events that coordinate the activities of the intestinal tract, pancreas, liver, and gallbladder.

Hormonal Responses to Food

The **endocrine system** controls the response of the various glands in the body and the release of hormones at the appropriate times.

One of the important factors under hormonal

control is the stomach acid environment. During the gastric phase, the hormone **gastrin** is secreted by G cells in the stomach in response to the presence of proteins. Gastrin stimulates the release of stomach acid, or hydrochloric acid (HCl) which aids in the digestion of the proteins. However, when the stomach is emptied, the acidic environment need not be maintained and a hormone called **somatostatin** stops the release of hydrochloric acid. This is controlled by a negative feedback mechanism.

In the duodenum, digestive secretions from the liver, pancreas, and gallbladder play an important role in digesting chyme during the intestinal phase. In order to neutralize the acidic chyme, a hormone called **secretin** stimulates the pancreas to produce alkaline bicarbonate solution and deliver it to the duodenum. Secretin acts in tandem with another hormone called **cholecystokinin** (CCK). Not only does CCK stimulate the pancreas to produce the requisite pancreatic juices, it also stimulates the gallbladder to release bile into the duodenum.

Link to Learning

Visit this website to learn more about the endocrine system. Review the text and watch the animation of how control is implemented in the endocrine system.

Another level of hormonal control occurs in response to the composition of food. Foods high in lipids take a long time to digest. A hormone called **gastric inhibitory peptide** is secreted by the small intestine to slow down the peristaltic movements of the intestine to allow fatty foods more time to be digested and absorbed.

Understanding the hormonal control of the digestive system is an important area of ongoing research. Scientists are exploring the role of each hormone in the digestive process and developing ways to target these hormones. Advances could lead to knowledge that may help to battle the obesity epidemic.

Section Summary

The brain and the endocrine system control digestive processes. The brain controls the responses of hunger and satiety. The endocrine system controls the release of hormones and enzymes required for digestion of food in the digestive tract.

Review Questions

Which hormone controls the release of bile from the gallbladder

1. pepsin
2. amylase
3. CCK
4. gastrin

C

Which hormone stops acid secretion in the stomach?

1. gastrin
2. somatostatin
3. gastric inhibitory peptide
4. CCK

B

In the famous conditioning experiment, Pavlov demonstrated that his dogs started drooling in response to a bell sounding. What part of the digestive process did he stimulate?

1. cephalic phase
2. gastric phase
3. intestinal phase
4. elimination phase

A

Critical Thinking Questions

Describe how hormones regulate digestion.

Hormones control the different digestive enzymes that are secreted in the stomach and the intestine during the process of digestion and absorption. For example, the hormone gastrin stimulates stomach acid secretion in response to food intake. The hormone somatostatin stops the release of stomach acid.

Describe one or more scenarios where loss of hormonal regulation of digestion can lead to diseases.

There are many cases where loss of hormonal regulation can lead to illnesses. For example, the bilirubin produced by the breakdown of red blood cells is converted to bile by the liver. When there is malfunction of this process, there is excess bilirubin in the blood and bile levels are low. As a result, the body struggles with

dealing with fatty food. This is why a patient suffering from jaundice is asked to eat a diet with almost zero fat.

A scientist is studying a model that has a mutation in the receptor for somatostatin that prevents hormone binding. How would this mutation affect the structure and function of the digestive system?

Somatostatin is the hormone that inhibits the release of HCl into the stomach lumen after the chyme has moved to the intestine. If the receptor for somatostatin is nonfunctional, somatostatin cannot signal to the stomach parietal cells to stop acid secretion. Thus, acid secretion will continue when there is no food present, and can cause damage to the stomach tissue. However, as long as the stomach remains intact the mutation should not slow digestion since acid will always be present in the stomach to digest any new boluses of food.

Glossary

cephalic phase
: first phase of digestion, controlled by the neural response to the stimulus provided by food

cholecystokinin
: hormone that stimulates the contraction of the gallbladder to release bile

endocrine system
: system that controls the response of the various glands in the body and the release of hormones at the appropriate times

gastric inhibitory peptide
: hormone secreted by the small intestine in the presence of fatty acids and sugars; it also inhibits acid production and peristalsis in order to slow down the rate at which food enters the small intestine

gastric phase
: digestive phase beginning once food enters the stomach; gastric acids and enzymes process the ingested materials

gastrin
: hormone which stimulates hydrochloric acid secretion in the stomach

intestinal phase
: third digestive phase; begins when chyme enters the small intestine triggering digestive secretions and controlling the rate of gastric emptying

secretin

hormone which stimulates sodium bicarbonate secretion in the small intestine

somatostatin
: hormone released to stop acid secretion when the stomach is empty

Introduction

class="introduction" An athlete’s nervous system is hard at work during the planning and execution of a movement as precise as a high jump. Parts of the nervous system are involved in determining how hard to push off and when to turn, as well as controlling the muscles throughout the body that make this complicated movement possible without knocking the bar down—all in just a few seconds. (credit: modification of work by Shane T. McCoy, U.S. Navy)

When you’re reading this book, your nervous system is performing several functions simultaneously. The visual system is processing what is seen on the page; the motor system controls the turn of the pages (or click of the mouse); the prefrontal cortex maintains attention. Even fundamental functions, like breathing and regulation of body temperature, are controlled by the nervous system. A nervous system is an organism’s control center: it processes sensory information from outside (and inside) the body and

controls all behaviors—from eating to sleeping to finding a mate.

Neurons and Glial Cells

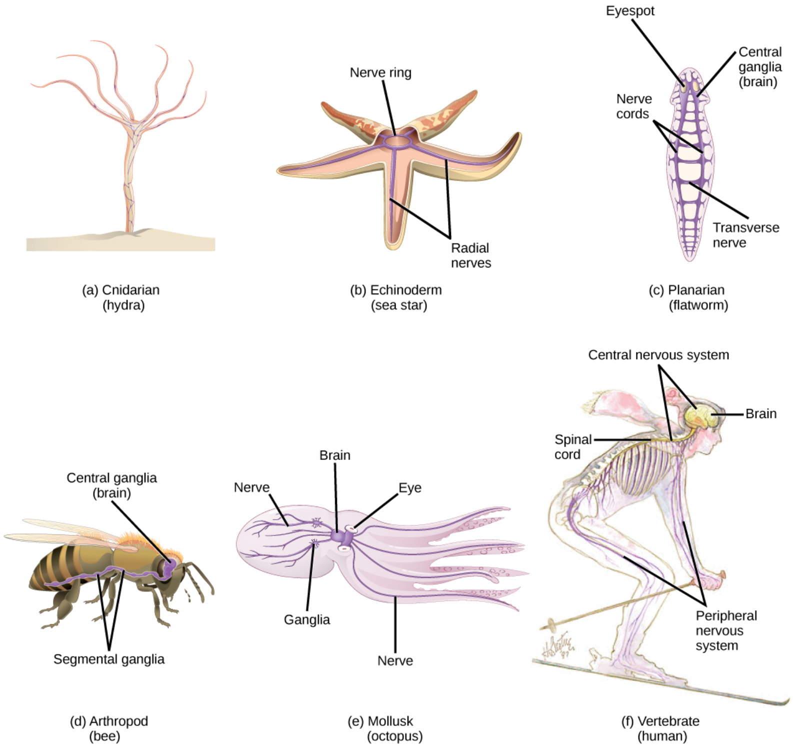
By the end of this section, you will be able to do the following:

- List and describe the functions of the structural components of a neuron
- List and describe the four main types of neurons
- Compare the functions of different types of glial cells

Nervous systems throughout the animal kingdom vary in structure and complexity, as illustrated by the variety of animals shown in [link]. Some organisms, like sea sponges, lack a true nervous system. Others, like jellyfish, lack a true brain and instead have a system of separate but connected nerve cells (neurons) called a “nerve net.” Echinoderms such as sea stars have nerve cells that are bundled into fibers called nerves. Flatworms of the phylum Platyhelminthes have both a central nervous system (CNS), made up of a small “brain” and two nerve cords, and a peripheral nervous system (PNS) containing a system of nerves that extend throughout the body. The insect nervous system is more complex but also fairly decentralized. It contains a brain, ventral nerve cord, and ganglia (clusters of connected neurons). These ganglia can control movements and behaviors without input from the brain. Octopi may have the most complicated of invertebrate nervous systems—

they have neurons that are organized in specialized lobes and eyes that are structurally similar to vertebrate species.

Nervous systems vary in structure and complexity. In (a) cnidarians, nerve cells form a decentralized nerve net. In (b) echinoderms, nerve cells are bundled into fibers called nerves. In animals exhibiting bilateral symmetry such as (c) planarians, neurons cluster into an anterior brain that processes information. In addition to a brain, (d) arthropods have clusters of nerve cell bodies, called peripheral ganglia, located along the ventral nerve cord. Mollusks such as squid and (e) octopi, which must hunt to survive, have complex brains containing millions of neurons. In (f) vertebrates, the brain and spinal cord comprise the central nervous system, while neurons extending into the rest of the body comprise the peripheral nervous system. (credit e: modification of work by Michael Vecchione, Clyde F.E. Roper, and Michael J. Sweeney, NOAA; credit f: modification of work by NIH)



Compared to invertebrates, vertebrate nervous systems are more complex, centralized, and specialized. While there is great diversity among different vertebrate nervous systems, they all share a basic structure: a CNS that contains a brain and spinal cord and a PNS made up of peripheral sensory and motor nerves. One interesting difference between the nervous systems of invertebrates and vertebrates is that the nerve cords of many invertebrates are located ventrally whereas the vertebrate spinal cords are located dorsally. There is debate among evolutionary biologists as to whether these different nervous system plans evolved separately or whether the invertebrate body

plan arrangement somehow “flipped” during the evolution of vertebrates.

Link to Learning
Watch this video of biologist Mark Kirschner discussing the “flipping” phenomenon of vertebrate evolution.
https://www.openstax.org/l/vertebrate_evol

The nervous system is made up of **neurons**, specialized cells that can receive and transmit chemical or electrical signals, and **glia**, cells that provide support functions for the neurons by playing an information processing role that is complementary to neurons. A neuron can be compared to an electrical wire—it transmits a signal from one place to another. Glia can be compared to the workers at the electric company who make sure wires go to the right places, maintain the wires, and take down wires that are broken. Although glia have been compared to workers, recent evidence suggests that they also usurp some of the signaling functions of neurons.

There is great diversity in the types of neurons and glia that are present in different parts of the nervous system. There are four major types of neurons, and

they share several important cellular components. There is great diversity in the size and shape of neurons throughout the nervous system. Examples include (a) a pyramidal cell from the cerebral cortex, (b) a Purkinje cell from the cerebellar cortex, and (c) olfactory cells from the olfactory epithelium and olfactory bulb. Neurons are broadly divided into four main types based on the number and placement of axons: (1) unipolar, (2) bipolar, (3) multipolar, and (4) pseudounipolar.

Neurons

The nervous system of the common laboratory fly, *Drosophila melanogaster*, contains around 100,000 neurons, the same number as a lobster. This number compares to 75 million in the mouse and 300 million in the octopus. A human brain contains around 86 billion neurons. Despite these very different numbers, the nervous systems of these animals control many of the same behaviors—from basic reflexes to more complicated behaviors like finding food and courting mates. The ability of neurons to communicate with each other as well as with other types of cells underlies all of these behaviors.

Most neurons share the same cellular components. But neurons are also highly specialized—different types of neurons have different sizes and shapes that relate to their functional roles.

Parts of a Neuron

Like other cells, each neuron has a cell body (or soma) that contains a nucleus, smooth and rough endoplasmic reticulum, Golgi apparatus, mitochondria, and other cellular components. Neurons also contain unique structures, illustrated in [link] for receiving and sending the electrical signals that make neuronal communication possible. **Dendrites** are tree-like structures that extend away from the cell body to receive messages from other neurons at specialized junctions called **synapses**. Although some neurons do not have any dendrites, some types of neurons have multiple dendrites. Dendrites can have small protrusions called dendritic spines, which further increase surface area for possible synaptic connections.

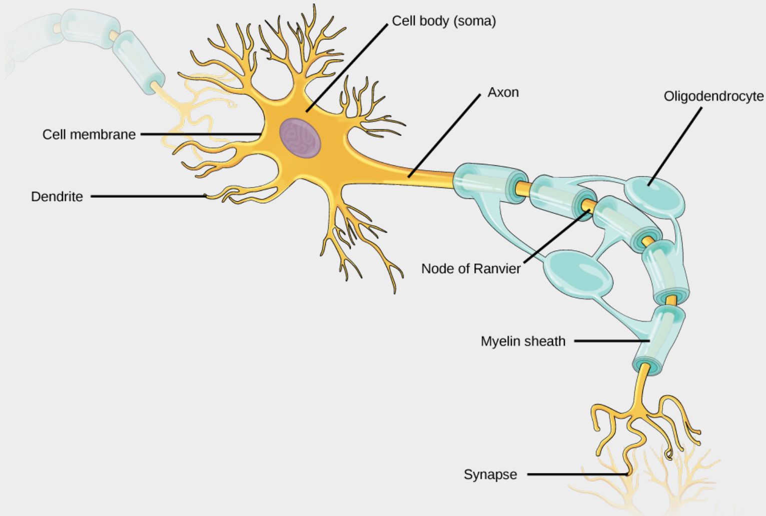
Once a signal is received by the dendrite, it then travels passively to the cell body. The cell body contains a specialized structure, the **axon hillock** that integrates signals from multiple synapses and serves as a junction between the cell body and an **axon**. An axon is a tube-like structure that propagates the integrated signal to specialized endings called **axon terminals**. These terminals in turn synapse on other neurons, muscle, or target organs. Chemicals released at axon terminals allow signals to be communicated to these other cells. Neurons usually have one or two axons, but some neurons, like amacrine cells in the retina, do not

contain any axons. Some axons are covered with **myelin**, which acts as an insulator to minimize dissipation of the electrical signal as it travels down the axon, greatly increasing the speed of conduction. This insulation is important as the axon from a human motor neuron can be as long as a meter—from the base of the spine to the toes. The myelin sheath is not actually part of the neuron. Myelin is produced by glial cells. Along the axon there are periodic gaps in the myelin sheath. These gaps are called **nodes of Ranvier** and are sites where the signal is “recharged” as it travels along the axon.

It is important to note that a single neuron does not act alone—neuronal communication depends on the connections that neurons make with one another (as well as with other cells, like muscle cells). Dendrites from a single neuron may receive synaptic contact from many other neurons. For example, dendrites from a Purkinje cell in the cerebellum are thought to receive contact from as many as 200,000 other neurons.

Visual Connection

Neurons contain organelles common to many other cells, such as a nucleus and mitochondria. They also have more specialized structures, including dendrites and axons.

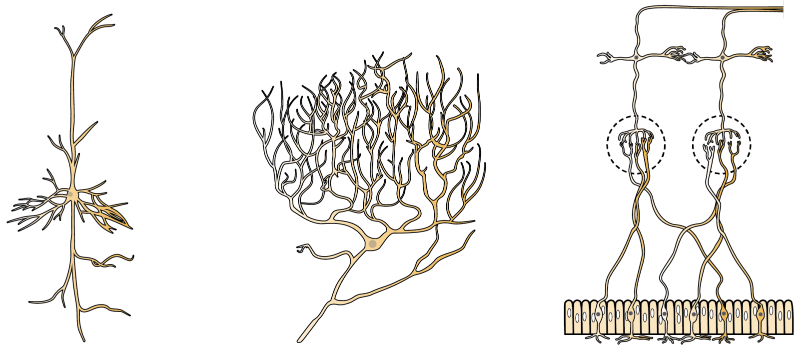


Which of the following statements is false?

1. The soma is the cell body of a nerve cell.
2. Myelin sheath provides an insulating layer to the dendrites.
3. Axons carry the signal from the soma to the target.
4. Dendrites carry the signal to the soma.

Types of Neurons

There are different types of neurons, and the functional role of a given neuron is intimately dependent on its structure. There is an amazing diversity of neuron shapes and sizes found in different parts of the nervous system (and across species), as illustrated by the neurons shown in [link].



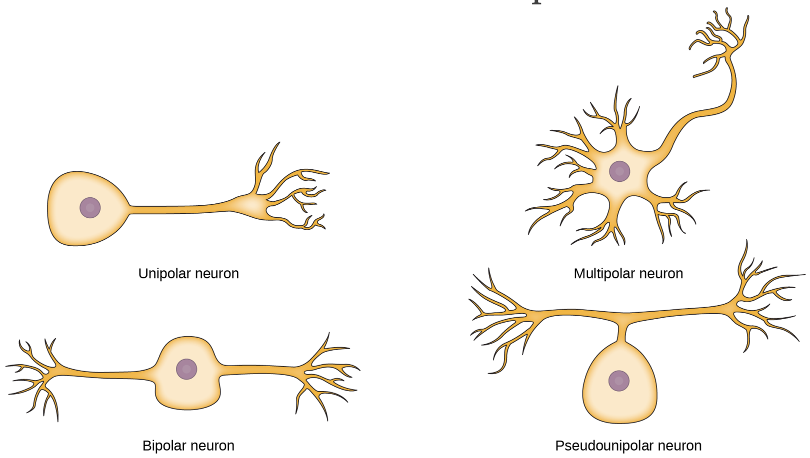
(a) Pyramidal cell of the cerebral cortex

(b) Purkinje cell of the cerebellar cortex

(c) Olfactory neurons

While there are many defined neuron cell subtypes, neurons are broadly divided into four basic types: unipolar, bipolar, multipolar, and pseudounipolar. [link] illustrates these four basic neuron types. Unipolar neurons have only one structure that extends away from the soma. These neurons are not found in vertebrates but are found in insects where they stimulate muscles or glands. A bipolar neuron has one axon and one dendrite extending from the soma. An example of a bipolar neuron is a retinal bipolar cell, which receives signals from photoreceptor cells that are sensitive to light and transmits these signals to ganglion cells that carry the signal to the brain. Multipolar neurons are the most common type of neuron. Each multipolar neuron contains one axon and multiple dendrites. Multipolar neurons can be found in the central nervous system (brain and spinal cord). An example of a multipolar neuron is a Purkinje cell in the cerebellum, which has many branching dendrites but only one axon. Pseudounipolar cells share characteristics with both unipolar and bipolar cells.

A pseudounipolar cell has a single process that extends from the soma, like a unipolar cell, but this process later branches into two distinct structures, like a bipolar cell. Most sensory neurons are pseudounipolar and have an axon that branches into two extensions: one connected to dendrites that receive sensory information and another that transmits this information to the spinal cord.



Everyday Connection

Neurogenesis

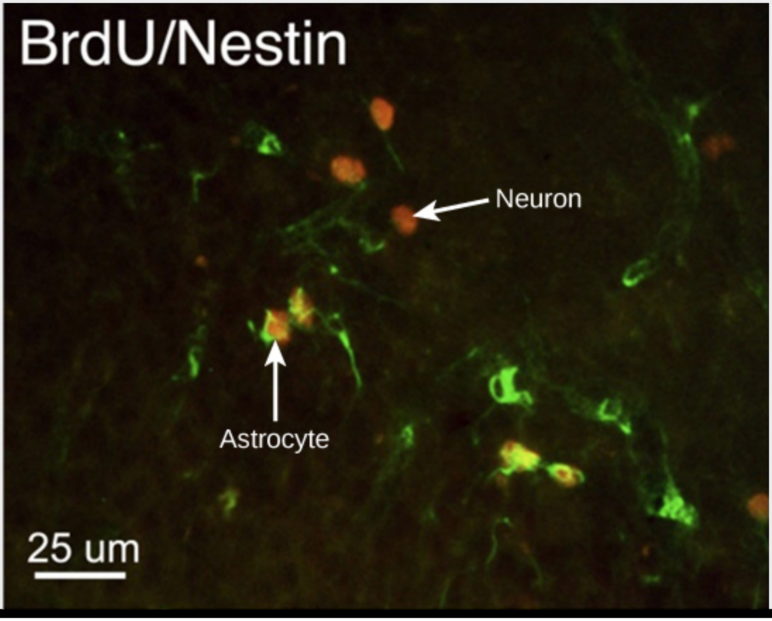
At one time, scientists believed that people were born with all the neurons they would ever have. Research performed during the last few decades indicates that neurogenesis, the birth of new neurons, continues into adulthood. Neurogenesis was first discovered in songbirds that produce new neurons while learning songs. For mammals, new neurons also play an important role in learning:

about 1000 new neurons develop in the hippocampus (a brain structure involved in learning and memory) each day. While most of the new neurons will die, researchers found that an increase in the number of surviving new neurons in the hippocampus correlated with how well rats learned a new task. Interestingly, both exercise and some antidepressant medications also promote neurogenesis in the hippocampus. Stress has the opposite effect. While neurogenesis is quite limited compared to regeneration in other tissues, research in this area may lead to new treatments for disorders such as Alzheimer's, stroke, and epilepsy.

How do scientists identify new neurons? A researcher can inject a compound called bromodeoxyuridine (BrdU) into the brain of an animal. While all cells will be exposed to BrdU, BrdU will only be incorporated into the DNA of newly generated cells that are in S phase. A technique called immunohistochemistry can be used to attach a fluorescent label to the incorporated BrdU, and a researcher can use fluorescent microscopy to visualize the presence of BrdU, and thus new neurons, in brain tissue. [link] is a micrograph which shows fluorescently labeled neurons in the hippocampus of a rat.

This micrograph shows fluorescently labeled new neurons in a rat hippocampus. Cells that are actively dividing have bromodoxyuridine (BrdU) incorporated into their DNA and are labeled in red. Cells that express glial fibrillary acidic protein

(GFAP) are labeled in green. Astrocytes, but not neurons, express GFAP. Thus, cells that are labeled both red and green are actively dividing astrocytes, whereas cells labeled red only are actively dividing neurons. (credit: modification of work by Dr. Maryam Faiz, et. al., University of Barcelona; scale-bar data from Matt Russell)



Link to Learning

This site contains more information about neurogenesis, including an interactive laboratory simulation and a video that explains how BrdU labels new cells.

Glial cells support neurons and maintain their environment. Glial cells of the (a) central nervous system include oligodendrocytes, astrocytes, ependymal cells, and microglial cells. Oligodendrocytes form the myelin sheath around axons. Astrocytes provide nutrients to neurons, maintain their extracellular environment, and provide structural support. Microglia scavenge pathogens and dead cells. Ependymal cells produce cerebrospinal fluid that cushions the neurons. Glial cells of the (b) peripheral nervous system include Schwann cells, which form the myelin sheath, and satellite cells, which provide nutrients and structural support to neurons. (a) Astrocytes and (b) oligodendrocytes are glial cells of the central nervous system. (credit a: modification of work by Uniformed Services University; credit b: modification of work by Jurjen Broeke; scale-bar data from Matt Russell)

Glia

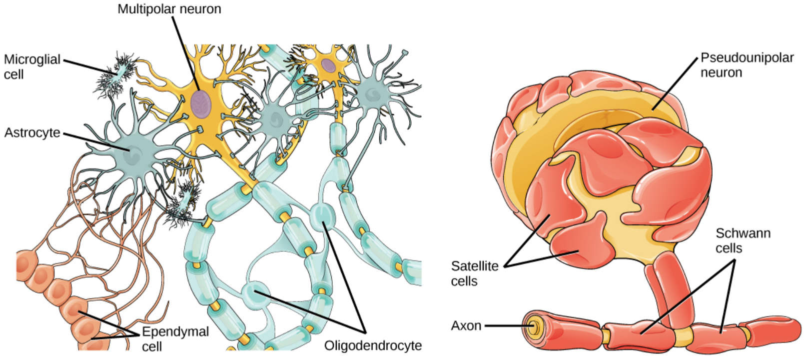
While glia are often thought of as the supporting cast of the nervous system, the number of glial cells in the brain actually outnumbers the number of neurons by a factor of ten. Neurons would be unable to function without the vital roles that are fulfilled by these glial cells. Glia guide developing neurons to their destinations, buffer ions and chemicals that would otherwise harm neurons, and provide myelin sheaths around axons. Scientists have recently

discovered that they also play a role in responding to nerve activity and modulating communication between nerve cells. When glia do not function properly, the result can be disastrous—most brain tumors are caused by mutations in glia.

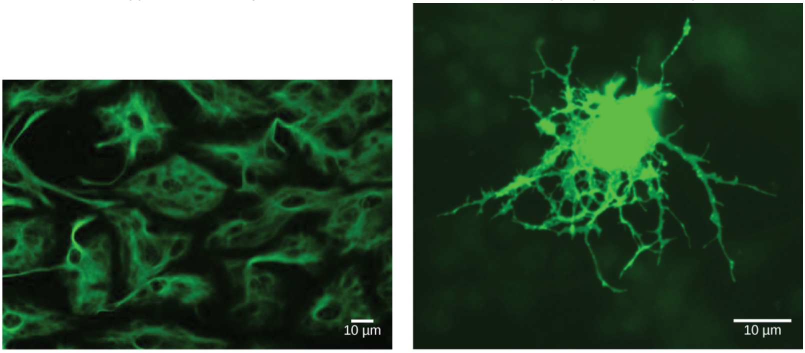
Types of Glia

There are several different types of glia with different functions, two of which are shown in [link]. **Astrocytes**, shown in [link]**a** make contact with both capillaries and neurons in the CNS. They provide nutrients and other substances to neurons, regulate the concentrations of ions and chemicals in the extracellular fluid, and provide structural support for synapses. Astrocytes also form the blood-brain barrier—a structure that blocks entrance of toxic substances into the brain. Astrocytes, in particular, have been shown through calcium imaging experiments to become active in response to nerve activity, transmit calcium waves between astrocytes, and modulate the activity of surrounding synapses. **Satellite glia** provide nutrients and structural support for neurons in the PNS. **Microglia** scavenge and degrade dead cells and protect the brain from invading microorganisms. **Oligodendrocytes**, shown in [link]**b** form myelin sheaths around axons in the CNS. One axon can be myelinated by several oligodendrocytes, and one oligodendrocyte can provide myelin for multiple neurons. This is

distinctive from the PNS where a single **Schwann cell** provides myelin for only one axon as the entire Schwann cell surrounds the axon. **Radial glia** serve as scaffolds for developing neurons as they migrate to their end destinations. **Ependymal** cells line fluid-filled ventricles of the brain and the central canal of the spinal cord. They are involved in the production of cerebrospinal fluid, which serves as a cushion for the brain, moves the fluid between the spinal cord and the brain, and is a component for the choroid plexus.



(a) Central nervous system
(b) Peripheral nervous system



(a) Astrocyte
(b) Oligodendrocyte

Section Summary

The nervous system is made up of neurons and glia. Neurons are specialized cells that are capable of sending electrical as well as chemical signals. Most neurons contain dendrites, which receive these signals, and axons that send signals to other neurons or tissues. There are four main types of neurons: unipolar, bipolar, multipolar, and pseudounipolar neurons. Glia are non-neuronal cells in the nervous system that support neuronal development and signaling. There are several types of glia that serve different functions.

Visual Connection Questions

[link] Which of the following statements is false?

1. The soma is the cell body of a nerve cell.
2. Myelin sheath provides an insulating layer to the dendrites.
3. Axons carry the signal from the soma to the target.
4. Dendrites carry the signal to the soma.

[link] B

Review Questions

Neurons contain ______, which can receive signals from other neurons.

1. axons
2. mitochondria
3. dendrites
4. Golgi bodies

C

A(n) ______ neuron has one axon and one dendrite extending directly from the cell body.

1. unipolar
2. bipolar
3. multipolar
4. pseudounipolar

B

Glia that provide myelin for neurons in the brain are called ______.

1. Schwann cells
2. oligodendrocytes

3. microglia
4. astrocytes

B

Meningitis is a viral or bacterial infection of the brain. Which cell type is the first to have its function disrupted during meningitis?

1. astrocytes
2. microglia
3. neurons
4. satellite glia

A

Critical Thinking Questions

How are neurons similar to other cells? How are they unique?

Neurons contain organelles common to all cells, such as a nucleus and mitochondria. They are unique because they contain dendrites, which can receive signals from other neurons, and

axons that can send these signals to other cells.

Multiple sclerosis causes demyelination of axons in the brain and spinal cord. Why is this problematic?

Myelin provides insulation for signals traveling along axons. Without myelin, signal transmission can slow down and degrade over time. This would slow down neuronal communication across the nervous system and affect all downstream functions.

Many neurons have only a single axon, but many terminals at the end of the axon. How does this end structure of the axon support its function?

A single axon means that a neuron can only send one signal at a time (one electrical impulse down the length of the axon). However, since the axon has multiple terminals it can send the signal to several other cells at once. This ensures that the signal is rapidly propagated to the rest of the body.

Glossary

astrocyte
: glial cell in the central nervous system that provide nutrients, extracellular buffering, and structural support for neurons; also makes up the blood-brain barrier

axon
: tube-like structure that propagates a signal from a neuron's cell body to axon terminals

axon hillock
: electrically sensitive structure on the cell body of a neuron that integrates signals from multiple neuronal connections

axon terminal
: structure on the end of an axon that can form a synapse with another neuron

dendrite
: structure that extends away from the cell body to receive messages from other neurons

ependymal
: cell that lines fluid-filled ventricles of the brain and the central canal of the spinal cord; involved in production of cerebrospinal fluid

glia
: (also, glial cells) cells that provide support functions for neurons

microglia
: glia that scavenge and degrade dead cells and protect the brain from invading microorganisms

myelin
: fatty substance produced by glia that insulates axons

neuron
: specialized cell that can receive and transmit electrical and chemical signals

nodes of Ranvier
: gaps in the myelin sheath where the signal is recharged

oligodendrocyte
: glial cell that myelinates central nervous system neuron axons

radial glia
: glia that serve as scaffolds for developing neurons as they migrate to their final destinations

satellite glia
: glial cell that provides nutrients and structural support for neurons in the peripheral nervous system

Schwann cell

glial cell that creates myelin sheath around a peripheral nervous system neuron axon

synapse
: junction between two neurons where neuronal signals are communicated

How Neurons Communicate

By the end of this section, you will be able to do the following:

- Describe the basis of the resting membrane potential
- Explain the stages of an action potential and how action potentials are propagated
- Explain the similarities and differences between chemical and electrical synapses
- Describe long-term potentiation and long-term depression

All functions performed by the nervous system—from a simple motor reflex to more advanced functions like making a memory or a decision—require neurons to communicate with one another. While humans use words and body language to communicate, neurons use electrical and chemical signals. Just like a person in a committee, one neuron usually receives and synthesizes messages from multiple other neurons before "making the decision" to send the message on to other neurons. Voltage-gated ion channels open in response to changes in membrane voltage. After activation, they become inactivated for a brief period and will no longer open in response to a signal. The (a) resting membrane potential is a result of different concentrations of Na^+ and K^+ ions inside and outside the cell. A nerve impulse causes Na^+ to enter the cell, resulting in (b) depolarization. At the

peak action potential, K+ channels open and the cell becomes (c) hyperpolarized. The action potential is conducted down the axon as the axon membrane depolarizes, then repolarizes. Nodes of Ranvier are gaps in myelin coverage along axons. Nodes contain voltage-gated K+ and Na+ channels. Action potentials travel down the axon by jumping from one node to the next.

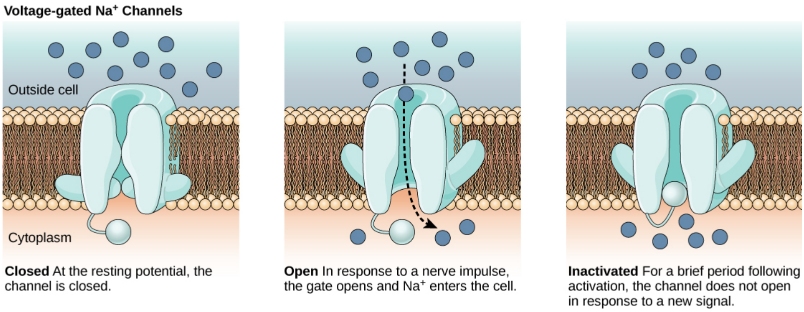
Nerve Impulse Transmission within a Neuron

For the nervous system to function, neurons must be able to send and receive signals. These signals are possible because each neuron has a charged cellular membrane (a voltage difference between the inside and the outside), and the charge of this membrane can change in response to neurotransmitter molecules released from other neurons and environmental stimuli. To understand how neurons communicate, one must first understand the basis of the baseline or 'resting' membrane charge.

Neuronal Charged Membranes

The lipid bilayer membrane that surrounds a neuron is impermeable to charged molecules or ions. To enter or exit the neuron, ions must pass through special proteins called ion channels that span the membrane. Ion channels have different

configurations: open, closed, and inactive, as illustrated in [link]. Some ion channels need to be activated in order to open and allow ions to pass into or out of the cell. These ion channels are sensitive to the environment and can change their shape accordingly. Ion channels that change their structure in response to voltage changes are called voltage-gated ion channels. Voltage-gated ion channels regulate the relative concentrations of different ions inside and outside the cell. The difference in total charge between the inside and outside of the cell is called the **membrane potential**.



Link to Learning

This video discusses the basis of the resting membrane potential.

https://www.openstax.org/l/resting_neuron

Resting Membrane Potential

A neuron at rest is negatively charged: the inside of a cell is approximately 70 millivolts more negative than the outside (−70 mV, note that this number varies by neuron type and by species). This voltage is called the resting membrane potential; it is caused by differences in the concentrations of ions inside and outside the cell. If the membrane were equally permeable to all ions, each type of ion would flow across the membrane and the system would reach equilibrium. Because ions cannot simply cross the membrane at will, there are different concentrations of several ions inside and outside the cell, as shown in [link]. The difference in the number of positively charged potassium ions (K^+) inside and outside the cell dominates the resting membrane potential ([link]). When the membrane is at rest, K^+ ions accumulate inside the cell due to a net movement with the concentration gradient. The negative resting membrane potential is created and maintained by increasing the concentration of cations outside the cell (in the extracellular fluid) relative to inside the cell (in the cytoplasm). The negative charge within the cell is created by the cell membrane being more permeable to potassium ion movement than sodium ion movement. In neurons, potassium ions are maintained at high concentrations within the cell while sodium ions are maintained at high concentrations outside of the cell. The cell possesses potassium and sodium leakage channels that allow the two cations to diffuse down their concentration gradient. However,

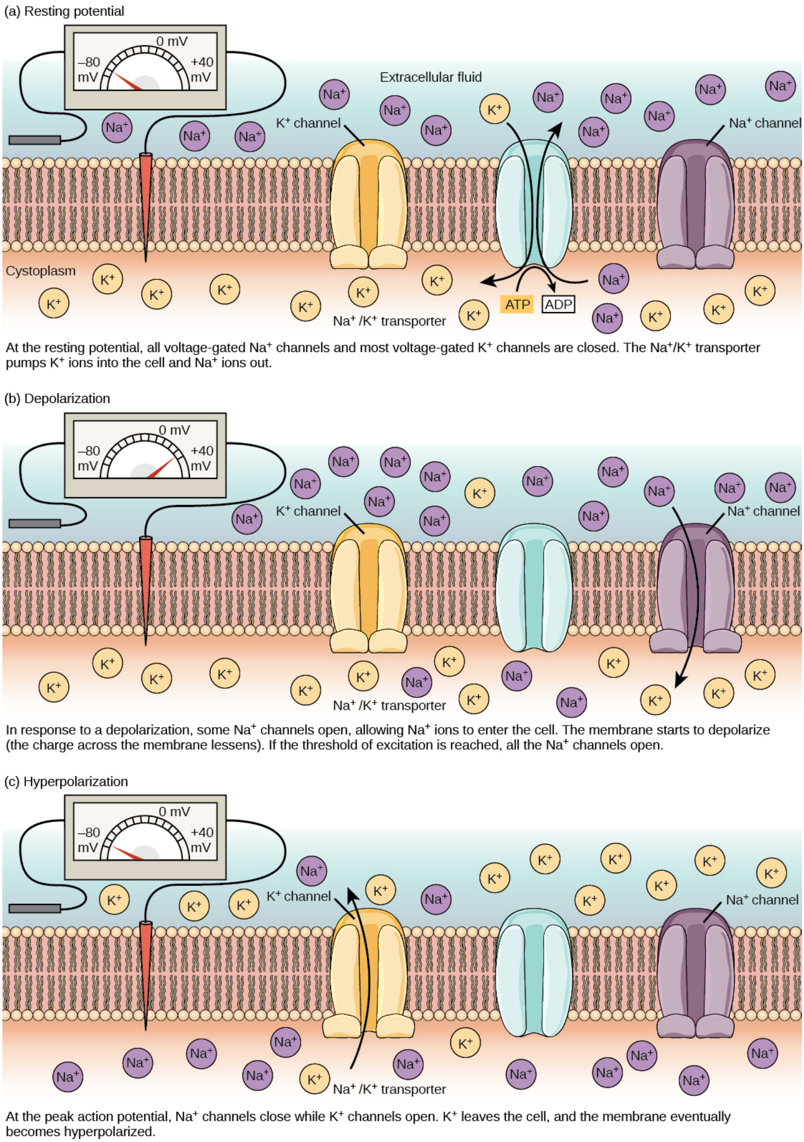
the neurons have far more potassium leakage channels than sodium leakage channels. Therefore, potassium diffuses out of the cell at a much faster rate than sodium leaks in. Because more cations are leaving the cell than are entering, this causes the interior of the cell to be negatively charged relative to the outside of the cell. The actions of the sodium potassium pump help to maintain the resting potential, once established. Recall that sodium potassium pumps brings two $K+$ ions into the cell while removing three $Na+$ ions per ATP consumed. As more cations are expelled from the cell than taken in, the inside of the cell remains negatively charged relative to the extracellular fluid. It should be noted that chloride ions ($Cl–$) tend to accumulate outside of the cell because they are repelled by negatively-charged proteins within the cytoplasm.

Ion Concentration Inside and Outside Neurons

Ion	Extracellular concentration (mM)	Intracellular concentration (mM)	Ratio outside/ inside
$Na+$	145	12	12

K+	4	155	0.026
Cl	120	4	30
Organic anions (A−)	—	100	

The resting membrane potential is a result of different concentrations inside and outside the cell.



Action Potential

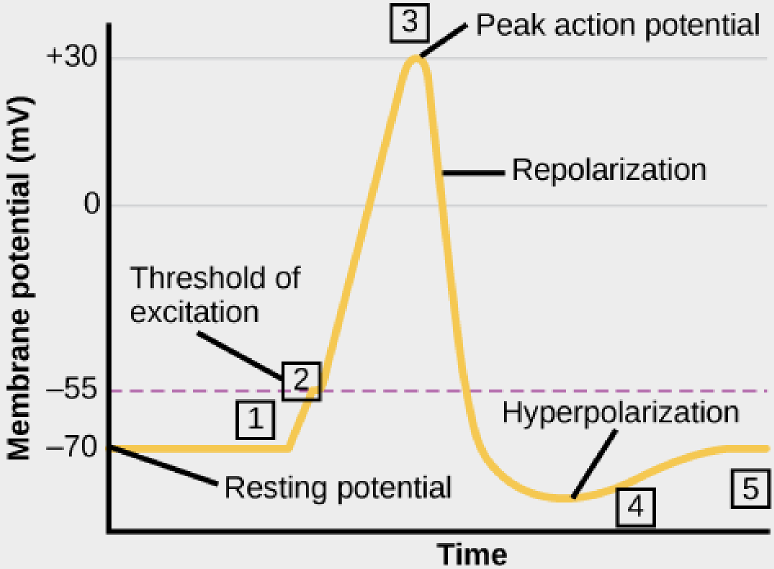
A neuron can receive input from other neurons and, if this input is strong enough, send the signal to downstream neurons. Transmission of a signal

between neurons is generally carried by a chemical called a neurotransmitter. Transmission of a signal within a neuron (from dendrite to axon terminal) is carried by a brief reversal of the resting membrane potential called an **action potential**. When neurotransmitter molecules bind to receptors located on a neuron's dendrites, ion channels open. At excitatory synapses, this opening allows positive ions to enter the neuron and results in **depolarization** of the membrane—a decrease in the difference in voltage between the inside and outside of the neuron. A stimulus from a sensory cell or another neuron depolarizes the target neuron to its threshold potential (-55 mV). Na^+ channels in the axon hillock open, allowing positive ions to enter the cell ([link] and [link]). Once the sodium channels open, the neuron completely depolarizes to a membrane potential of about $+40$ mV. Action potentials are considered an "all-or nothing" event, in that, once the threshold potential is reached, the neuron always completely depolarizes. Once depolarization is complete, the cell must now "reset" its membrane voltage back to the resting potential. To accomplish this, the Na^+ channels close and cannot be opened. This begins the neuron's **refractory period**, in which it cannot produce another action potential because its sodium channels will not open. At the same time, voltage-gated K^+ channels open, allowing K^+ to leave the cell. As K^+ ions leave the cell, the membrane potential once again becomes negative. The

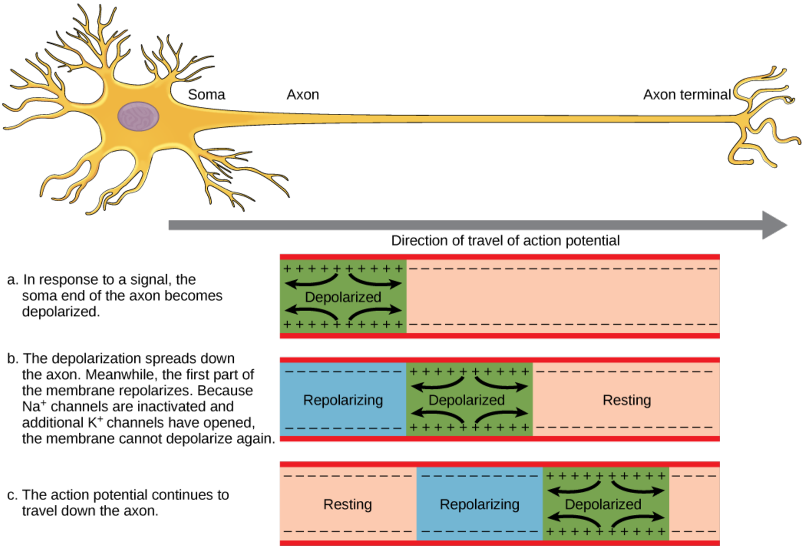
diffusion of K+ out of the cell actually **hyperpolarizes** the cell, in that the membrane potential becomes more negative than the cell's normal resting potential. At this point, the sodium channels will return to their resting state, meaning they are ready to open again if the membrane potential again exceeds the threshold potential. Eventually the extra K+ ions diffuse out of the cell through the potassium leakage channels, bringing the cell from its hyperpolarized state, back to its resting membrane potential.

Visual Connection

The formation of an action potential can be divided into five steps: (1) A stimulus from a sensory cell or another neuron causes the target cell to depolarize toward the threshold potential. (2) If the threshold of excitation is reached, all Na+ channels open and the membrane depolarizes. (3) At the peak action potential, K+ channels open and K+ begins to leave the cell. At the same time, Na+ channels close. (4) The membrane becomes hyperpolarized as K+ ions continue to leave the cell. The hyperpolarized membrane is in a refractory period and cannot fire. (5) The K+ channels close and the Na+/K+ transporter restores the resting potential.



Potassium channel blockers, such as amiodarone and procainamide, which are used to treat abnormal electrical activity in the heart, called cardiac dysrhythmia, impede the movement of $K+$ through voltage-gated $K+$ channels. Which part of the action potential would you expect potassium channels to affect?



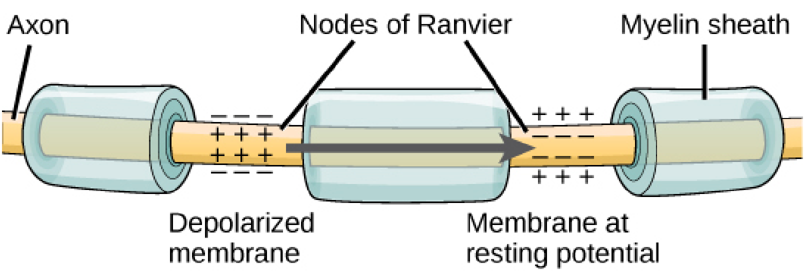
Link to Learning

This video presents an overview of action potential.

Myelin and the Propagation of the Action Potential

For an action potential to communicate information to another neuron, it must travel along the axon and reach the axon terminals where it can initiate neurotransmitter release. The speed of conduction of an action potential along an axon is influenced by both the diameter of the axon and the axon's resistance to current leak. Myelin acts as an insulator that prevents current from leaving the

axon; this increases the speed of action potential conduction. In demyelinating diseases like multiple sclerosis, action potential conduction slows because current leaks from previously insulated axon areas. The nodes of Ranvier, illustrated in [link] are gaps in the myelin sheath along the axon. These unmyelinated spaces are about one micrometer long and contain voltage-gated Na^+ and K^+ channels. Flow of ions through these channels, particularly the Na^+ channels, regenerates the action potential over and over again along the axon. This 'jumping' of the action potential from one node to the next is called **saltatory conduction**. If nodes of Ranvier were not present along an axon, the action potential would propagate very slowly since Na^+ and K^+ channels would have to continuously regenerate action potentials at every point along the axon instead of at specific points. Nodes of Ranvier also save energy for the neuron since the channels only need to be present at the nodes and not along the entire axon.



This pseudocolored image taken with a scanning electron microscope shows an axon terminal that was broken open to reveal synaptic vesicles (blue and orange) inside the neuron. (credit: modification of work by Tina Carvalho, NIH-NIGMS; scale-bar

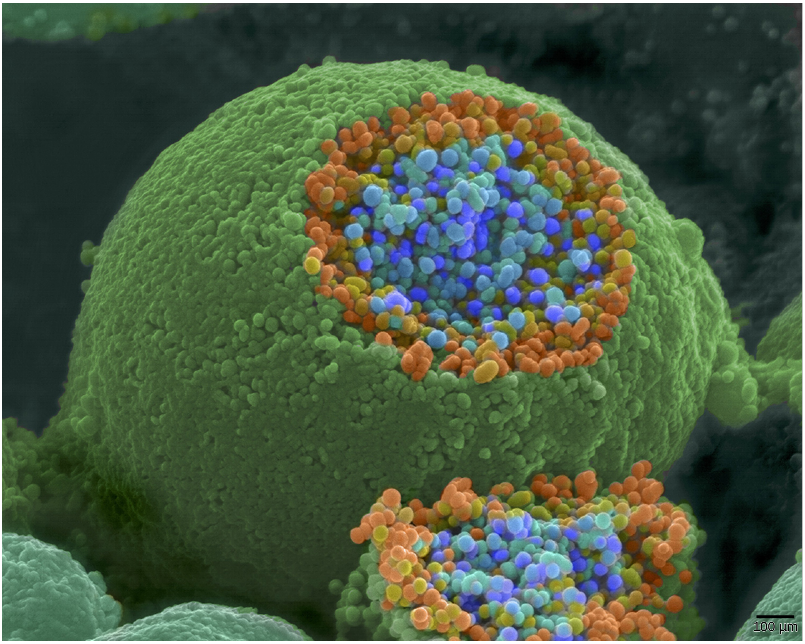
data from Matt Russell) Communication at chemical synapses requires release of neurotransmitters. When the presynaptic membrane is depolarized, voltage-gated Ca^{2+} channels open and allow Ca^{2+} to enter the cell. The calcium entry causes synaptic vesicles to fuse with the membrane and release neurotransmitter molecules into the synaptic cleft. The neurotransmitter diffuses across the synaptic cleft and binds to ligand-gated ion channels in the postsynaptic membrane, resulting in a localized depolarization or hyperpolarization of the postsynaptic neuron.

Synaptic Transmission

The synapse or "gap" is the place where information is transmitted from one neuron to another. Synapses usually form between axon terminals and dendritic spines, but this is not universally true. There are also axon-to-axon, dendrite-to-dendrite, and axon-to-cell body synapses. The neuron transmitting the signal is called the presynaptic neuron, and the neuron receiving the signal is called the postsynaptic neuron. Note that these designations are relative to a particular synapse—most neurons are both presynaptic and postsynaptic. There are two types of synapses: chemical and electrical.

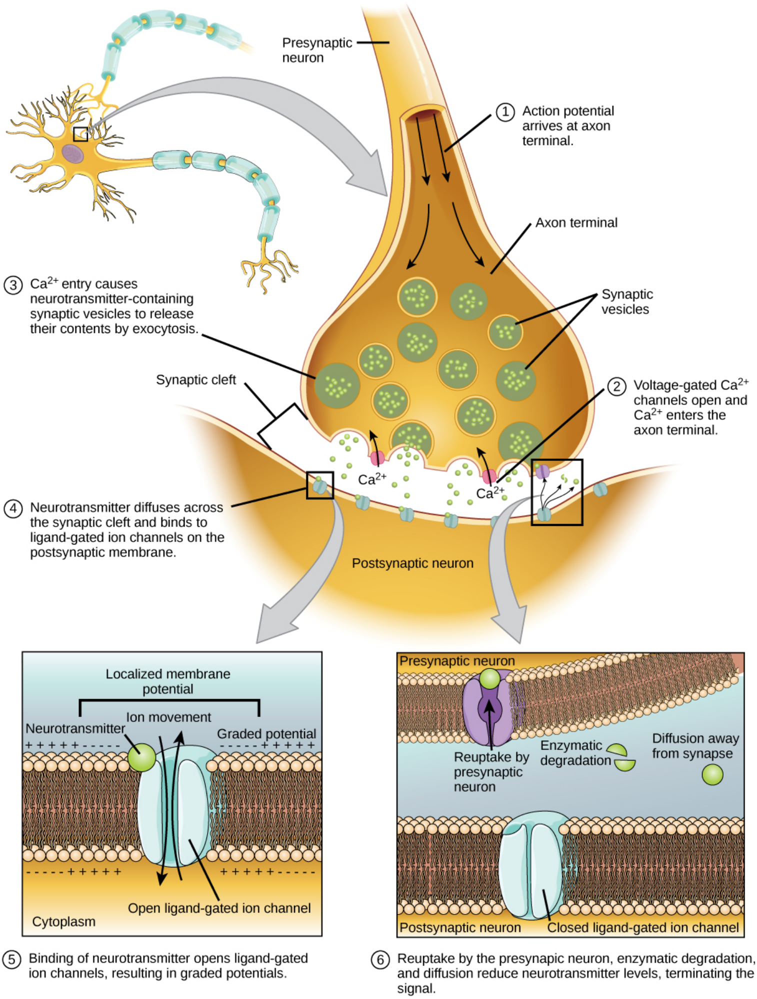
Chemical Synapse

When an action potential reaches the axon terminal it depolarizes the membrane and opens voltage-gated Na^+ channels. Na^+ ions enter the cell, further depolarizing the presynaptic membrane. This depolarization causes voltage-gated Ca^{2+} channels to open. Calcium ions entering the cell initiate a signaling cascade that causes small membrane-bound vesicles, called **synaptic vesicles**, containing neurotransmitter molecules to fuse with the presynaptic membrane. Synaptic vesicles are shown in [link], which is an image from a scanning electron microscope.



Fusion of a vesicle with the presynaptic membrane causes neurotransmitter to be released into the **synaptic cleft**, the extracellular space between the presynaptic and postsynaptic membranes, as

illustrated in [link]. The neurotransmitter diffuses across the synaptic cleft and binds to receptor proteins on the postsynaptic membrane.



The binding of a specific neurotransmitter causes particular ion channels, in this case ligand-gated channels, on the postsynaptic membrane to open. Neurotransmitters can either have excitatory or inhibitory effects on the postsynaptic membrane, as

detailed in [link]. For example, when acetylcholine is released at the synapse between a nerve and muscle (called the neuromuscular junction) by a presynaptic neuron, it causes postsynaptic Na^+ channels to open. Na^+ enters the postsynaptic cell and causes the postsynaptic membrane to depolarize. This depolarization is called an **excitatory postsynaptic potential (EPSP)** and makes the postsynaptic neuron more likely to fire an action potential. Release of neurotransmitter at inhibitory synapses causes **inhibitory postsynaptic potentials (IPSPs)**, a hyperpolarization of the presynaptic membrane. For example, when the neurotransmitter GABA (gamma-aminobutyric acid) is released from a presynaptic neuron, it binds to and opens Cl^- channels. Cl^- ions enter the cell and hyperpolarizes the membrane, making the neuron less likely to fire an action potential.

Once neurotransmission has occurred, the neurotransmitter must be removed from the synaptic cleft so the postsynaptic membrane can “reset” and be ready to receive another signal. This can be accomplished in three ways: the neurotransmitter can diffuse away from the synaptic cleft, it can be degraded by enzymes in the synaptic cleft, or it can be recycled (sometimes called reuptake) by the presynaptic neuron. Several drugs act at this step of neurotransmission. For example, some drugs that are given to Alzheimer’s patients work by inhibiting acetylcholinesterase, the enzyme

that degrades acetylcholine. This inhibition of the enzyme essentially increases neurotransmission at synapses that release acetylcholine. Once released, the acetylcholine stays in the cleft and can continually bind and unbind to postsynaptic receptors.

Neurotransmitter Function and Location

Neurotransmitter	Example	Location
Acetylcholine		CNS and/or PNS
Biogenic amine	Dopamine, serotonin, norepinephrine	CNS and/or PNS
Amino acid	Glycine, glutamate, aspartate, gamma aminobutyric acid	CNS
Neuropeptide	Substance P, endorphins	CNS and/or PNS

Electrical Synapse

While electrical synapses are fewer in number than chemical synapses, they are found in all nervous systems and play important and unique roles. The mode of neurotransmission in electrical synapses is quite different from that in chemical synapses. In an electrical synapse, the presynaptic and postsynaptic membranes are very close together and are actually physically connected by channel proteins forming gap junctions. Gap junctions allow current to pass directly from one cell to the next. In addition to the ions that carry this current, other molecules, such as ATP, can diffuse through the large gap junction pores.

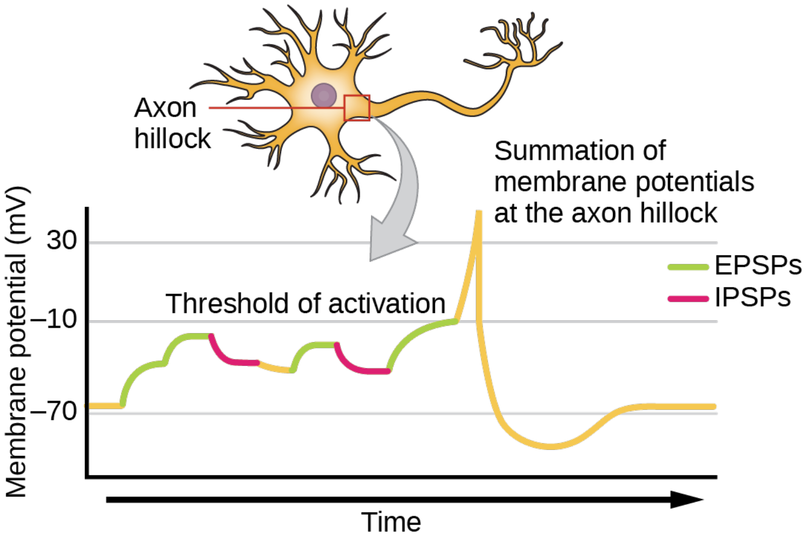
There are key differences between chemical and electrical synapses. Because chemical synapses depend on the release of neurotransmitter molecules from synaptic vesicles to pass on their signal, there is an approximately one millisecond delay between when the axon potential reaches the presynaptic terminal and when the neurotransmitter leads to opening of postsynaptic ion channels. Additionally, this signaling is unidirectional. Signaling in electrical synapses, in contrast, is virtually instantaneous (which is important for synapses involved in key reflexes), and some electrical synapses are bidirectional. Electrical synapses are also more reliable as they are less likely to be blocked, and they are important for synchronizing the electrical activity of a group of neurons. For example, electrical synapses in the thalamus are

thought to regulate slow-wave sleep, and disruption of these synapses can cause seizures.

A single neuron can receive both excitatory and inhibitory inputs from multiple neurons, resulting in local membrane depolarization (EPSP input) and hyperpolarization (IPSP input). All these inputs are added together at the axon hillock. If the EPSPs are strong enough to overcome the IPSPs and reach the threshold of excitation, the neuron will fire.

Signal Summation

Sometimes a single EPSP is strong enough to induce an action potential in the postsynaptic neuron, but often multiple presynaptic inputs must create EPSPs around the same time for the postsynaptic neuron to be sufficiently depolarized to fire an action potential. This process is called **summation** and occurs at the axon hillock, as illustrated in [link]. Additionally, one neuron often has inputs from many presynaptic neurons—some excitatory and some inhibitory—so IPSPs can cancel out EPSPs and vice versa. It is the net change in postsynaptic membrane voltage that determines whether the postsynaptic cell has reached its threshold of excitation needed to fire an action potential. Together, synaptic summation and the threshold for excitation act as a filter so that random “noise” in the system is not transmitted as important information.



Everyday Connection

Brain-computer interface

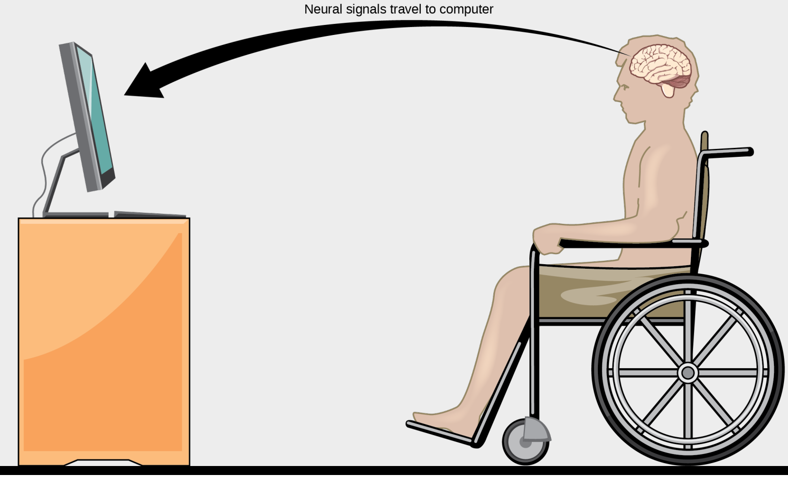
Amyotrophic lateral sclerosis (ALS, also called Lou Gehrig's Disease) is a neurological disease characterized by the degeneration of the motor neurons that control voluntary movements. The disease begins with muscle weakening and lack of coordination and eventually destroys the neurons that control speech, breathing, and swallowing; in the end, the disease can lead to paralysis. At that point, patients require assistance from machines to be able to breathe and to communicate. Several special technologies have been developed to allow "locked-in" patients to communicate with the rest of the world. One technology, for example, allows patients to type out sentences by twitching their

cheek. These sentences can then be read aloud by a computer.

A relatively new line of research for helping paralyzed patients, including those with ALS, to communicate and retain a degree of self-sufficiency is called brain-computer interface (BCI) technology and is illustrated in [link]. This technology sounds like something out of science fiction: it allows paralyzed patients to control a computer using only their thoughts. There are several forms of BCI. Some forms use EEG recordings from electrodes taped onto the skull. These recordings contain information from large populations of neurons that can be decoded by a computer. Other forms of BCI require the implantation of an array of electrodes smaller than a postage stamp in the arm and hand area of the motor cortex. This form of BCI, while more invasive, is very powerful as each electrode can record actual action potentials from one or more neurons. These signals are then sent to a computer, which has been trained to decode the signal and feed it to a tool—such as a cursor on a computer screen. This means that a patient with ALS can use e-mail, read the Internet, and communicate with others by thinking of moving his or her hand or arm (even though the paralyzed patient cannot make that bodily movement). Recent advances have allowed a paralyzed locked-in patient who suffered a stroke 15 years ago to control a robotic arm and even to feed herself coffee using BCI technology.

Despite the amazing advancements in BCI technology, it also has limitations. The technology can require many hours of training and long periods of intense concentration for the patient; it can also require brain surgery to implant the devices.

With brain-computer interface technology, neural signals from a paralyzed patient are collected, decoded, and then fed to a tool, such as a computer, a wheelchair, or a robotic arm.



Link to Learning

Watch this video in which a paralyzed woman uses a brain-controlled robotic arm to bring a drink to her mouth, among other images of brain-computer interface technology in action.

https://www.openstax.org/l/paralyzation

Calcium entry through postsynaptic NMDA receptors can initiate two different forms of synaptic plasticity: long-term potentiation (LTP) and long-term depression (LTD). LTP arises when a single synapse is repeatedly stimulated. This stimulation causes a calcium- and CaMKII-dependent cellular cascade, which results in the insertion of more AMPA receptors into the postsynaptic membrane. The next time glutamate is released from the presynaptic cell, it will bind to both NMDA and the newly inserted AMPA receptors, thus depolarizing the membrane more efficiently. LTD occurs when few glutamate molecules bind to NMDA receptors at a synapse (due to a low firing rate of the presynaptic neuron). The calcium that does flow through NMDA receptors initiates a different calcineurin and protein phosphatase 1-dependent cascade, which results in the endocytosis of AMPA receptors. This makes the postsynaptic neuron less responsive to glutamate released from the presynaptic neuron.

Synaptic Plasticity

Synapses are not static structures. They can be weakened or strengthened. They can be broken, and new synapses can be made. Synaptic plasticity allows for these changes, which are all needed for a functioning nervous system. In fact, synaptic plasticity is the basis of learning and memory. Two processes in particular, long-term potentiation (LTP)

and long-term depression (LTD) are important forms of synaptic plasticity that occur in synapses in the hippocampus, a brain region that is involved in storing memories.

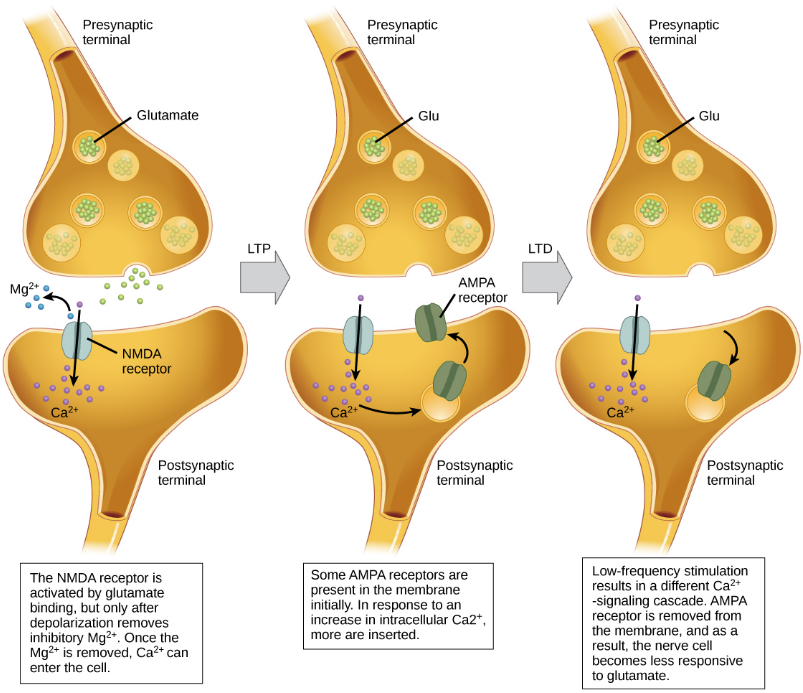
Long-term Potentiation (LTP)

Long-term potentiation (LTP) is a persistent strengthening of a synaptic connection. LTP is based on the Hebbian principle: cells that fire together wire together. There are various mechanisms, none fully understood, behind the synaptic strengthening seen with LTP. One known mechanism involves a type of postsynaptic glutamate receptor, called NMDA (N-Methyl-D-aspartate) receptors, shown in [link]. These receptors are normally blocked by magnesium ions; however, when the postsynaptic neuron is depolarized by multiple presynaptic inputs in quick succession (either from one neuron or multiple neurons), the magnesium ions are forced out allowing Ca ions to pass into the postsynaptic cell. Next, Ca^{2+} ions entering the cell initiate a signaling cascade that causes a different type of glutamate receptor, called AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) receptors, to be inserted into the postsynaptic membrane, since activated AMPA receptors allow positive ions to enter the cell. So, the next time glutamate is released from the presynaptic membrane, it will have a larger excitatory effect (EPSP) on the postsynaptic cell because the binding

of glutamate to these AMPA receptors will allow more positive ions into the cell. The insertion of additional AMPA receptors strengthens the synapse and means that the postsynaptic neuron is more likely to fire in response to presynaptic neurotransmitter release. Some drugs of abuse co-opt the LTP pathway, and this synaptic strengthening can lead to addiction.

Long-term Depression (LTD)

Long-term depression (LTD) is essentially the reverse of LTP: it is a long-term weakening of a synaptic connection. One mechanism known to cause LTD also involves AMPA receptors. In this situation, calcium that enters through NMDA receptors initiates a different signaling cascade, which results in the removal of AMPA receptors from the postsynaptic membrane, as illustrated in [link]. The decrease in AMPA receptors in the membrane makes the postsynaptic neuron less responsive to glutamate released from the presynaptic neuron. While it may seem counterintuitive, LTD may be just as important for learning and memory as LTP. The weakening and pruning of unused synapses allows for unimportant connections to be lost and makes the synapses that have undergone LTP that much stronger by comparison.



Section Summary

Neurons have charged membranes because there are different concentrations of ions inside and outside of the cell. Voltage-gated ion channels control the movement of ions into and out of a neuron. When a neuronal membrane is depolarized to at least the threshold of excitation, an action potential is fired. The action potential is then propagated along a myelinated axon to the axon terminals. In a chemical synapse, the action potential causes release of neurotransmitter molecules into the synaptic cleft. Through binding to postsynaptic receptors, the

neurotransmitter can cause excitatory or inhibitory postsynaptic potentials by depolarizing or hyperpolarizing, respectively, the postsynaptic membrane. In electrical synapses, the action potential is directly communicated to the postsynaptic cell through gap junctions—large channel proteins that connect the pre-and postsynaptic membranes. Synapses are not static structures and can be strengthened and weakened. Two mechanisms of synaptic plasticity are long-term potentiation and long-term depression.

Visual Connection Questions

[link] Potassium channel blockers, such as amiodarone and procainamide, which are used to treat abnormal electrical activity in the heart, called cardiac dysrhythmia, impede the movement of $K+$ through voltage-gated $K+$ channels. Which part of the action potential would you expect potassium channels to affect?

[link] Potassium channel blockers slow the repolarization phase, but have no effect on depolarization.

Review Questions

For a neuron to fire an action potential, its membrane must reach _____.

1. hyperpolarization
2. the threshold of excitation
3. the refractory period
4. inhibitory postsynaptic potential

B

After an action potential, the opening of additional voltage-gated _____ channels and the inactivation of sodium channels, cause the membrane to return to its resting membrane potential.

1. sodium
2. potassium
3. calcium
4. chloride

B

What is the term for protein channels that connect two neurons at an electrical synapse?

1. synaptic vesicles
2. voltage-gated ion channels
3. gap junction protein
4. sodium-potassium exchange pumps

C

Which of the following molecules is **not** involved in the maintenance of the resting membrane potential?

1. potassium cations
2. ATP
3. voltage-gated ion channels
4. calcium cations

D

Critical Thinking Questions

How does myelin aid propagation of an action potential along an axon? How do the nodes of Ranvier help this process?

Myelin prevents the leak of current from the

axon. Nodes of Ranvier allow the action potential to be regenerated at specific points along the axon. They also save energy for the cell since voltage-gated ion channels and sodium-potassium transporters are not needed along myelinated portions of the axon.

What are the main steps in chemical neurotransmission?

An action potential travels along an axon until it depolarizes the membrane at an axon terminal. Depolarization of the membrane causes voltage-gated Ca^{2+} channels to open and Ca^{2+} to enter the cell. The intracellular calcium influx causes synaptic vesicles containing neurotransmitter to fuse with the presynaptic membrane. The neurotransmitter diffuses across the synaptic cleft and binds to receptors on the postsynaptic membrane. Depending on the specific neurotransmitter and postsynaptic receptor, this action can cause positive (excitatory postsynaptic potential) or negative (inhibitory postsynaptic potential) ions to enter the cell.

Describe how long-term potentiation can lead to a nicotine addiction.

Long-term potentiation describes the process whereby exposure to a stimulus increases the likelihood that a neuron will depolarize in response to that stimulus in the future. Nicotine exposure causes long-term potentiation of neurons in the amygdala, and activates reward centers of the brain. As nicotine exposure continues, long-term potentiation reinforces the activation of the reward pathways in response to nicotine consumption.

Glossary

action potential
: self-propagating momentary change in the electrical potential of a neuron (or muscle) membrane

depolarization
: change in the membrane potential to a less negative value

excitatory postsynaptic potential (EPSP)
: depolarization of a postsynaptic membrane caused by neurotransmitter molecules released from a presynaptic cell

hyperpolarization
: change in the membrane potential to a more negative value

inhibitory postsynaptic potential (IPSP)
: hyperpolarization of a postsynaptic membrane caused by neurotransmitter molecules released from a presynaptic cell

long-term depression (LTD)
: prolonged decrease in synaptic coupling between a pre- and postsynaptic cell

long-term potentiation (LTP)
: prolonged increase in synaptic coupling between a pre-and postsynaptic cell

membrane potential
: difference in electrical potential between the inside and outside of a cell

refractory period
: period after an action potential when it is more difficult or impossible for an action potential to be fired; caused by inactivation of sodium channels and activation of additional potassium channels of the membrane

saltatory conduction
: “jumping” of an action potential along an axon from one node of Ranvier to the next

summation
: process of multiple presynaptic inputs creating EPSPs around the same time for the postsynaptic neuron to be sufficiently

depolarized to fire an action potential

synaptic cleft
: space between the presynaptic and postsynaptic membranes

synaptic vesicle
: spherical structure that contains a neurotransmitter

threshold of excitation
: level of depolarization needed for an action potential to fire

The Central Nervous System

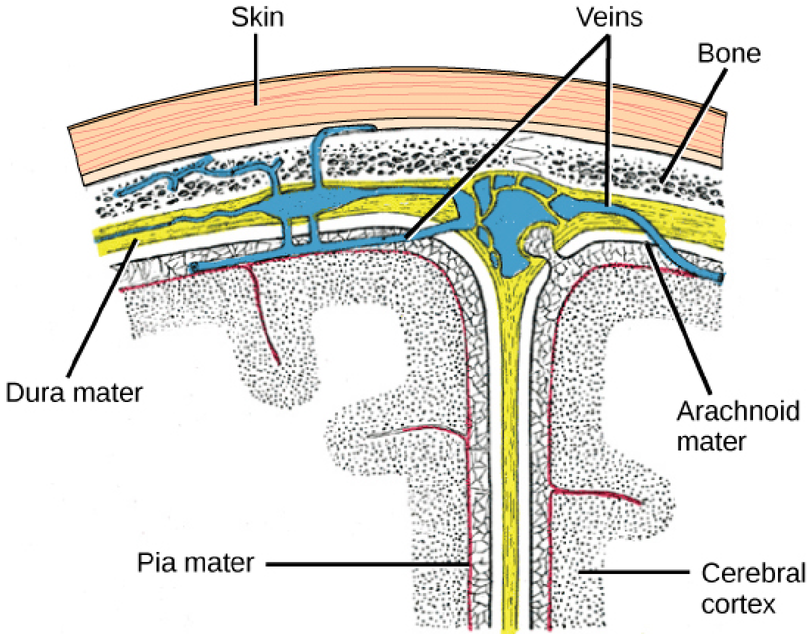
By the end of this section, you will be able to do the following:

- Identify the spinal cord, cerebral lobes, and other brain areas on a diagram of the brain
- Describe the basic functions of the spinal cord, cerebral lobes, and other brain areas

The central nervous system (CNS) is made up of the brain, a part of which is shown in [link] and spinal cord and is covered with three layers of protective coverings called **meninges** (from the Greek word for membrane). The outermost layer is the **dura mater** (Latin for “hard mother”). As the Latin suggests, the primary function for this thick layer is to protect the brain and spinal cord. The dura mater also contains vein-like structures that carry blood from the brain back to the heart. The middle layer is the web-like **arachnoid mater**. The last layer is the **pia mater** (Latin for “soft mother”), which directly contacts and covers the brain and spinal cord like plastic wrap. The space between the arachnoid and pia maters is filled with **cerebrospinal fluid (CSF)**. CSF is produced by a tissue called **choroid plexus** in fluid-filled compartments in the CNS called **ventricles**. The brain floats in CSF, which acts as a cushion and shock absorber and makes the brain neutrally buoyant. CSF also functions to circulate chemical substances throughout the brain and into the spinal cord.

The entire brain contains only about 8.5 tablespoons of CSF, but CSF is constantly produced in the ventricles. This creates a problem when a ventricle is blocked—the CSF builds up and creates swelling and the brain is pushed against the skull. This swelling condition is called hydrocephalus (“water head”) and can cause seizures, cognitive problems, and even death if a shunt is not inserted to remove the fluid and pressure.

The cerebral cortex is covered by three layers of meninges: the dura, arachnoid, and pia maters. (credit: modification of work by Gray’s Anatomy)



These illustrations show the (a) coronal and (b) sagittal sections of the human brain. The human cerebral cortex includes the frontal, parietal, temporal, and occipital lobes. Different parts of the motor cortex control different muscle groups.

Muscle groups that are neighbors in the body are generally controlled by neighboring regions of the motor cortex as well. For example, the neurons that control finger movement are near the neurons that control hand movement. The limbic system regulates emotion and other behaviors. It includes parts of the cerebral cortex located near the center of the brain, including the cingulate gyrus and the hippocampus as well as the thalamus, hypothalamus, and amygdala.

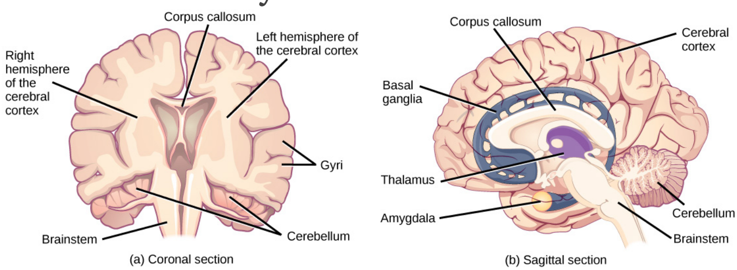
Brain

The brain is the part of the central nervous system that is contained in the cranial cavity of the skull. It includes the cerebral cortex, limbic system, basal ganglia, thalamus, hypothalamus, and cerebellum. There are three different ways that a brain can be sectioned in order to view internal structures: a sagittal section cuts the brain left to right, as shown in [link]**b**, a coronal section cuts the brain front to back, as shown in [link]**a**, and a horizontal section cuts the brain top to bottom.

Cerebral Cortex

The outermost part of the brain is a thick piece of nervous system tissue called the **cerebral cortex**, which is folded into hills called **gyri** (singular: gyrus) and valleys called **sulci** (singular: sulcus).

The cortex is made up of two hemispheres—right and left—which are separated by a large sulcus. A thick fiber bundle called the **corpus callosum** (Latin: "tough body") connects the two hemispheres and allows information to be passed from one side to the other. Although there are some brain functions that are localized more to one hemisphere than the other, the functions of the two hemispheres are largely redundant. In fact, sometimes (very rarely) an entire hemisphere is removed to treat severe epilepsy. While patients do suffer some deficits following the surgery, they can have surprisingly few problems, especially when the surgery is performed on children who have very immature nervous systems.



In other surgeries to treat severe epilepsy, the corpus callosum is cut instead of removing an entire hemisphere. This causes a condition called split-brain, which gives insights into unique functions of the two hemispheres. For example, when an object is presented to patients' left visual field, they may be unable to verbally name the object (and may claim to not have seen an object at all). This is because the visual input from the left visual field crosses and enters the right hemisphere and cannot

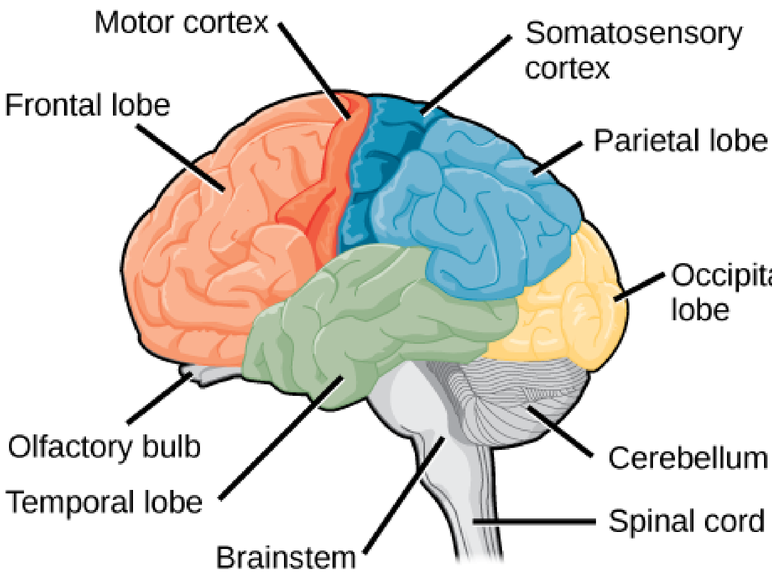
then signal to the speech center, which generally is found in the left side of the brain. Remarkably, if a split-brain patient is asked to pick up a specific object out of a group of objects with the left hand, the patient will be able to do so but will still be unable to vocally identify it.

Link to Learning

See this website to learn more about split-brain patients and to play a game where you can model the split-brain experiments yourself.

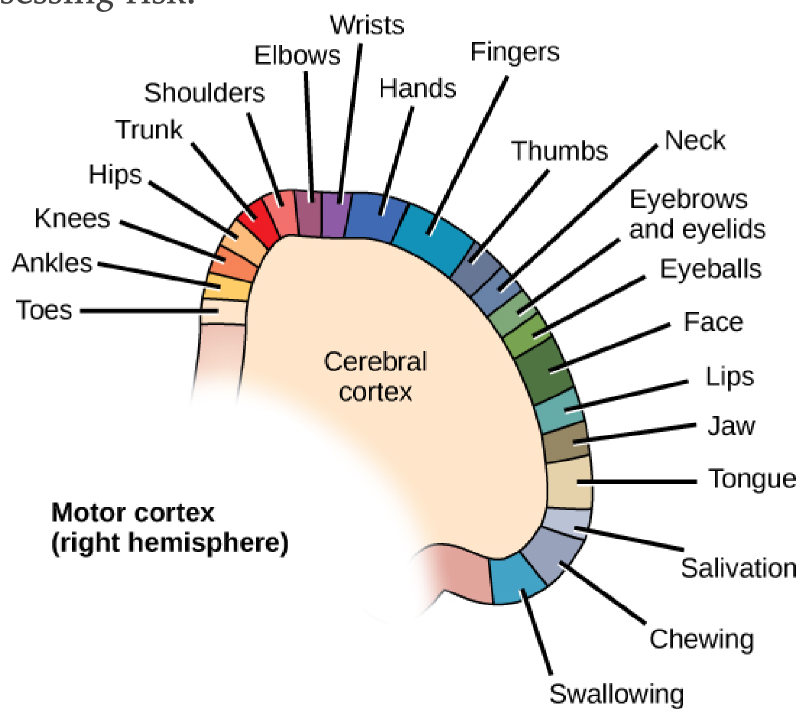
Each cortical hemisphere contains regions called lobes that are involved in different functions. Scientists use various techniques to determine what brain areas are involved in different functions: they examine patients who have had injuries or diseases that affect specific areas and see how those areas are related to functional deficits. They also conduct animal studies where they stimulate brain areas and see if there are any behavioral changes. They use a technique called transmagnetic stimulation (TMS) to temporarily deactivate specific parts of the cortex using strong magnets placed outside the head; and they use functional magnetic resonance imaging (fMRI) to look at changes in oxygenated blood flow in particular brain regions that correlate with

specific behavioral tasks. These techniques, and others, have given great insight into the functions of different brain regions but have also showed that any given brain area can be involved in more than one behavior or process, and any given behavior or process generally involves neurons in multiple brain areas. That being said, each hemisphere of the mammalian cerebral cortex can be broken down into four functionally and spatially defined lobes: frontal, parietal, temporal, and occipital. [link] illustrates these four lobes of the human cerebral cortex.



The **frontal lobe** is located at the front of the brain, over the eyes. This lobe contains the olfactory bulb, which processes smells. The frontal lobe also contains the motor cortex, which is important for

planning and implementing movement. Areas within the motor cortex map to different muscle groups, and there is some organization to this map, as shown in [link]. For example, the neurons that control movement of the fingers are next to the neurons that control movement of the hand. Neurons in the frontal lobe also control cognitive functions like maintaining attention, speech, and decision-making. Studies of humans who have damaged their frontal lobes show that parts of this area are involved in personality, socialization, and assessing risk.



The **parietal lobe** is located at the top of the brain. Neurons in the parietal lobe are involved in speech and also reading. Two of the parietal lobe's main

functions are processing **somatosensation**—touch sensations like pressure, pain, heat, cold—and processing **proprioception**—the sense of how parts of the body are oriented in space. The parietal lobe contains a somatosensory map of the body similar to the motor cortex.

The **occipital lobe** is located at the back of the brain. It is primarily involved in vision—seeing, recognizing, and identifying the visual world.

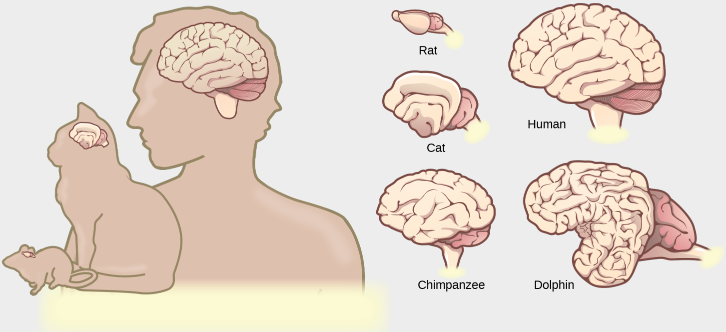
The **temporal lobe** is located at the base of the brain by your ears and is primarily involved in processing and interpreting sounds. It also contains the **hippocampus** (Greek for "seahorse")—a structure that processes memory formation. The hippocampus is illustrated in [link]. The role of the hippocampus in memory was partially determined by studying one famous epileptic patient, HM, who had both sides of his hippocampus removed in an attempt to cure his epilepsy. His seizures went away, but he could no longer form new memories (although he could remember some facts from before his surgery and could learn new motor tasks).

Evolution Connection

Cerebral Cortex

Compared to other vertebrates, mammals have exceptionally large brains for their body size. An

entire alligator's brain, for example, would fill about one and a half teaspoons. This increase in brain to body size ratio is especially pronounced in apes, whales, and dolphins. While this increase in overall brain size doubtlessly played a role in the evolution of complex behaviors unique to mammals, it does not tell the whole story. Scientists have found a relationship between the relatively high surface area of the cortex and the intelligence and complex social behaviors exhibited by some mammals. This increased surface area is due, in part, to increased folding of the cortical sheet (more sulci and gyri). For example, a rat cortex is very smooth with very few sulci and gyri. Cat and sheep cortices have more sulci and gyri. Chimps, humans, and dolphins have even more. Mammals have larger brain-to-body ratios than other vertebrates. Within mammals, increased cortical folding and surface area is correlated with complex behavior.



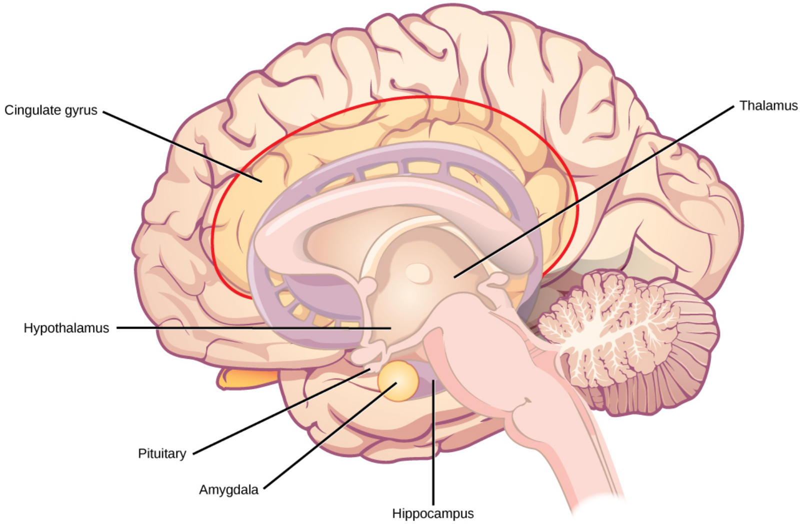
Basal Ganglia

Interconnected brain areas called the **basal ganglia** (or **basal nuclei**), shown in [link]**b**, play important roles in movement control and posture. Damage to the basal ganglia, as in Parkinson's disease, leads to motor impairments like a shuffling gait when walking. The basal ganglia also regulate motivation. For example, when a wasp sting led to bilateral basal ganglia damage in a 25-year-old businessman, he began to spend all his days in bed and showed no interest in anything or anybody. But when he was externally stimulated—as when someone asked to play a card game with him—he was able to function normally. Interestingly, he and other similar patients do not report feeling bored or frustrated by their state.

Thalamus

The **thalamus** (Greek for "inner chamber"), illustrated in [link], acts as a gateway to and from the cortex. It receives sensory and motor inputs from the body and also receives feedback from the cortex. This feedback mechanism can modulate conscious awareness of sensory and motor inputs depending on the attention and arousal state of the animal. The thalamus helps regulate consciousness, arousal, and sleep states. A rare genetic disorder called fatal familial insomnia causes the degeneration of thalamic neurons and glia. This

disorder prevents affected patients from being able to sleep, among other symptoms, and is eventually fatal.



Hypothalamus

Below the thalamus is the **hypothalamus**, shown in [link]. The hypothalamus controls the endocrine system by sending signals to the pituitary gland, a pea-sized endocrine gland that releases several different hormones that affect other glands as well as other cells. This relationship means that the hypothalamus regulates important behaviors that are controlled by these hormones. The hypothalamus is the body’s thermostat—it makes sure key functions like food and water intake, energy expenditure, and body temperature are kept at appropriate levels. Neurons within the hypothalamus also regulate circadian rhythms,

sometimes called sleep cycles.

Limbic System

The **limbic system** is a connected set of structures that regulates emotion, as well as behaviors related to fear and motivation. It plays a role in memory formation and includes parts of the thalamus and hypothalamus as well as the hippocampus. One important structure within the limbic system is a temporal lobe structure called the **amygdala** (Greek for “almond”), illustrated in [link]. The two amygdala are important both for the sensation of fear and for recognizing fearful faces. The **cingulate gyrus** helps regulate emotions and pain.

Cerebellum

The **cerebellum** (Latin for “little brain”), shown in [link], sits at the base of the brain on top of the brainstem. The cerebellum controls balance and aids in coordinating movement and learning new motor tasks.

Brainstem

The **brainstem**, illustrated in [link], connects the rest of the brain with the spinal cord. It consists of the midbrain, medulla oblongata, and the pons. Motor and sensory neurons extend through the

brainstem allowing for the relay of signals between the brain and spinal cord. Ascending neural pathways cross in this section of the brain allowing the left hemisphere of the cerebrum to control the right side of the body and vice versa. The brainstem coordinates motor control signals sent from the brain to the body. The brainstem controls several important functions of the body including alertness, arousal, breathing, blood pressure, digestion, heart rate, swallowing, walking, and sensory and motor information integration.

A cross-section of the spinal cord shows gray matter (containing cell bodies and interneurons) and white matter (containing axons).

Spinal Cord

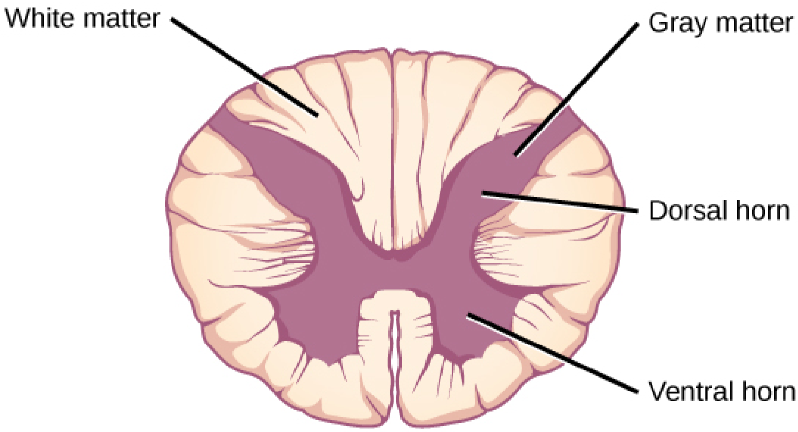
Connecting to the brainstem and extending down the body through the spinal column is the **spinal cord**, shown in [link]. The spinal cord is a thick bundle of nerve tissue that carries information about the body to the brain and from the brain to the body. The spinal cord is contained within the bones of the vertebrate column but is able to communicate signals to and from the body through its connections with spinal nerves (part of the peripheral nervous system). A cross-section of the spinal cord looks like a white oval containing a gray butterfly-shape, as illustrated in [link]. Myelinated axons make up the "white matter" and neuron and glial cell bodies make up the "gray matter." Gray matter is also

composed of interneurons, which connect two neurons each located in different parts of the body. Axons and cell bodies in the dorsal (facing the back of the animal) spinal cord convey mostly sensory information from the body to the brain. Axons and cell bodies in the ventral (facing the front of the animal) spinal cord primarily transmit signals controlling movement from the brain to the body.

The spinal cord also controls motor reflexes. These reflexes are quick, unconscious movements—like automatically removing a hand from a hot object. Reflexes are so fast because they involve local synaptic connections. For example, the knee reflex that a doctor tests during a routine physical is controlled by a single synapse between a sensory neuron and a motor neuron. While a reflex may only require the involvement of one or two synapses, synapses with interneurons in the spinal column transmit information to the brain to convey what happened (the knee jerked, or the hand was hot).

In the United States, there around 10,000 spinal cord injuries each year. Because the spinal cord is the information superhighway connecting the brain with the body, damage to the spinal cord can lead to paralysis. The extent of the paralysis depends on the location of the injury along the spinal cord and whether the spinal cord was completely severed. For example, if the spinal cord is damaged at the level of the neck, it can cause paralysis from the neck

down, whereas damage to the spinal column further down may limit paralysis to the legs. Spinal cord injuries are notoriously difficult to treat because spinal nerves do not regenerate, although ongoing research suggests that stem cell transplants may be able to act as a bridge to reconnect severed nerves. Researchers are also looking at ways to prevent the inflammation that worsens nerve damage after injury. One such treatment is to pump the body with cold saline to induce hypothermia. This cooling can prevent swelling and other processes that are thought to worsen spinal cord injuries.



Section Summary

The vertebrate central nervous system contains the brain and the spinal cord, which are covered and protected by three meninges. The brain contains structurally and functionally defined regions. In mammals, these include the cortex (which can be

broken down into four primary functional lobes: frontal, temporal, occipital, and parietal), basal ganglia, thalamus, hypothalamus, limbic system, cerebellum, and brainstem—although structures in some of these designations overlap. While functions may be primarily localized to one structure in the brain, most complex functions, like language and sleep, involve neurons in multiple brain regions. The spinal cord is the information superhighway that connects the brain with the rest of the body through its connections with peripheral nerves. It transmits sensory and motor input and also controls motor reflexes.

Review Questions

The ______ lobe contains the visual cortex.

1. frontal
2. parietal
3. temporal
4. occipital

D

The ______ connects the two cerebral hemispheres.

1. limbic system
2. corpus callosum
3. cerebellum
4. pituitary

B

Neurons in the ______ control motor reflexes.

1. thalamus
2. spinal cord
3. parietal lobe
4. hippocampus

B

Phineas Gage was a 19th century railroad worker who survived an accident that drove a large iron rod through his head. If the injury resulted in him becoming temperamental and capricious what part of his brain was damaged?

1. frontal lobe
2. hippocampus
3. parietal lobe
4. temporal lobe

A

Critical Thinking Questions

What methods can be used to determine the function of a particular brain region?

To determine the function of a specific brain area, scientists can look at patients who have damage in that brain area and see what symptoms they exhibit. Researchers can disable the brain structure temporarily using transcranial magnetic stimulation. They can disable or remove the area in an animal model. fMRI can be used to correlate specific functions with increased blood flow to brain regions.

What are the main functions of the spinal cord?

The spinal cord transmits sensory information from the body to the brain and motor commands from the brain to the body through its connections with peripheral nerves. It also controls motor reflexes.

Alzheimer's disease involves three of the four lobes of the brain. Identify one of the involved lobes and describe the lobe's symptoms associated with the disease.

Potential answers:

1. Frontal lobe. Alzheimer's patients experience changes in personality, judgment, and behavior.;
2. Parietal lobe. Alzheimer's patients experience difficulties with recalling and using language as disease progresses.;
3. Temporal lobe. The hippocampus is one of the main areas of the brain affected in Alzheimer's disease. Patients lose the ability to make new memories and access memories.

Glossary

amygdala
: structure within the limbic system that processes fear

arachnoid mater
: spiderweb-like middle layer of the meninges that cover the central nervous system

basal ganglia
: interconnected collections of cells in the brain that are involved in movement and motivation; also known as basal nuclei

basal nuclei
: see basal ganglia

brainstem
: portion of the brain that connects with the spinal cord; controls basic nervous system functions like breathing, heart rate, and swallowing

cerebellum
: brain structure involved in posture, motor coordination, and learning new motor actions

cerebral cortex
: outermost sheet of brain tissue; involved in many higher-order functions

choroid plexus
: spongy tissue within ventricles that produces cerebrospinal fluid

cingulate gyrus
: helps regulate emotions and pain; thought to directly drive the body's conscious response to unpleasant experiences

corpus callosum

thick fiber bundle that connects the cerebral hemispheres

cerebrospinal fluid (CSF)
: clear liquid that surrounds the brain and spinal cord and fills the ventricles and central canal; acts as a shock absorber and circulates material throughout the brain and spinal cord

dura mater
: tough outermost layer that covers the central nervous system

frontal lobe
: part of the cerebral cortex that contains the motor cortex and areas involved in planning, attention, and language

gyrus
: (plural: gyri) ridged protrusions in the cortex

hippocampus
: brain structure in the temporal lobe involved in processing memories

hypothalamus
: brain structure that controls hormone release and body homeostasis

limbic system
: connected brain areas that process emotion and motivation

meninge
: membrane that covers and protects the central nervous system

occipital lobe
: part of the cerebral cortex that contains visual cortex and processes visual stimuli

parietal lobe
: part of the cerebral cortex involved in processing touch and the sense of the body in space

pia mater
: thin membrane layer directly covering the brain and spinal cord

proprioception
: sense about how parts of the body are oriented in space

somatosensation
: sense of touch

spinal cord
: thick fiber bundle that connects the brain with peripheral nerves; transmits sensory and motor information; contains neurons that control motor reflexes

sulcus
: (plural: sulci) indents or "valleys" in the

cortex

temporal lobe
: part of the cerebral cortex that processes auditory input; parts of the temporal lobe are involved in speech, memory, and emotion processing

thalamus
: brain area that relays sensory information to the cortex

ventricle
: cavity within brain that contains cerebrospinal fluid

The Peripheral Nervous System

By the end of this section, you will be able to do the following:

- Describe the organization and functions of the sympathetic and parasympathetic nervous systems
- Describe the organization and function of the sensory-somatic nervous system

The peripheral nervous system (PNS) is the connection between the central nervous system and the rest of the body. The CNS is like the power plant of the nervous system. It creates the signals that control the functions of the body. The PNS is like the wires that go to individual houses. Without those "wires," the signals produced by the CNS could not control the body (and the CNS would not be able to receive sensory information from the body either).

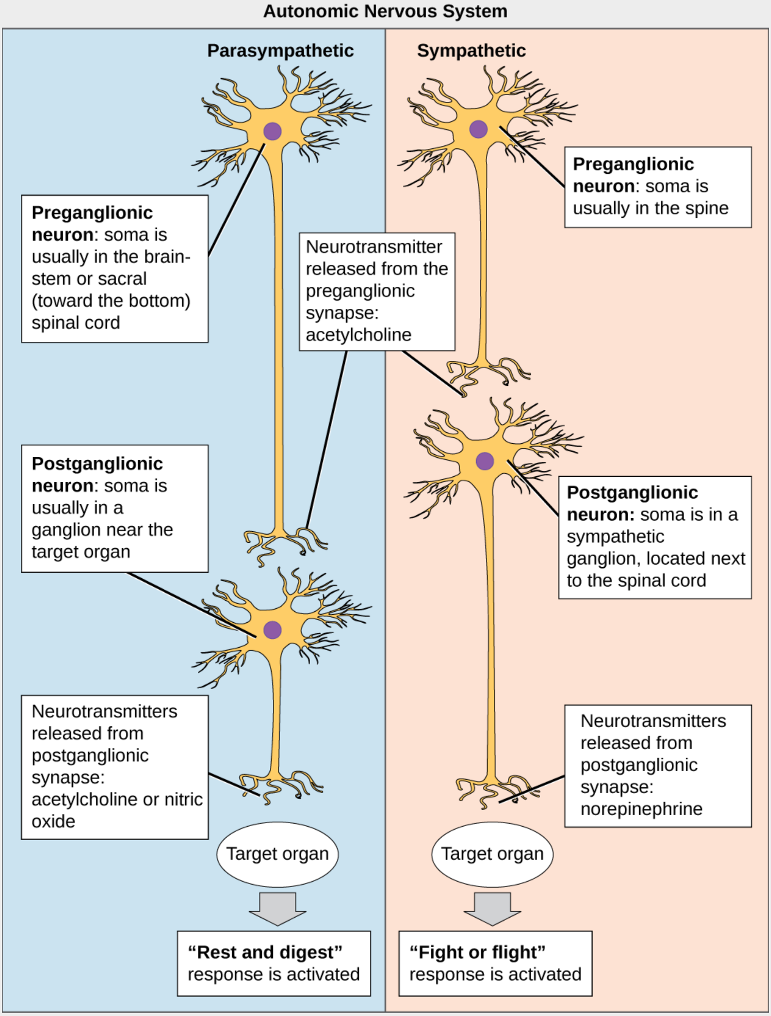
The PNS can be broken down into the **autonomic nervous system**, which controls bodily functions without conscious control, and the **sensory-somatic nervous system**, which transmits sensory information from the skin, muscles, and sensory organs to the CNS and sends motor commands from the CNS to the muscles.

The sympathetic and parasympathetic nervous systems often have opposing effects on target organs.

Autonomic Nervous System

Visual Connection

In the autonomic nervous system, a preganglionic neuron of the CNS synapses with a postganglionic neuron of the PNS. The postganglionic neuron, in turn, acts on a target organ. Autonomic responses are mediated by the sympathetic and the parasympathetic systems, which are antagonistic to one another. The sympathetic system activates the "fight or flight" response, while the parasympathetic system activates the "rest and digest" response.



Which of the following statements is false?

1. The parasympathetic pathway is responsible for resting the body, while the sympathetic pathway is responsible for preparing for an emergency.
2. Most preganglionic neurons in the sympathetic pathway originate in the spinal cord.
3. Slowing of the heartbeat is a parasympathetic

response.

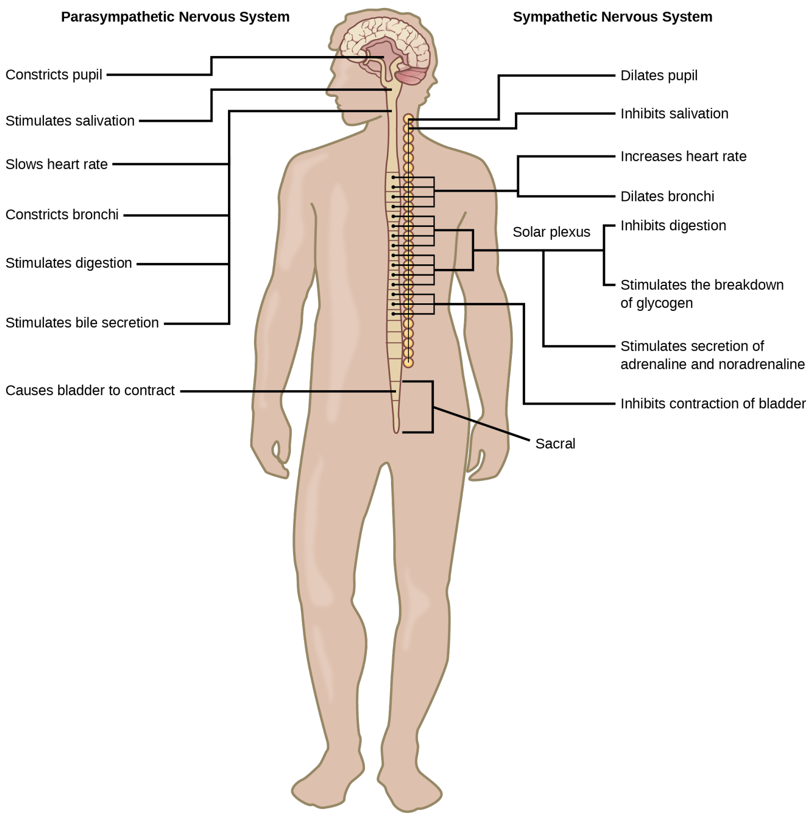
4. Parasympathetic neurons are responsible for releasing norepinephrine on the target organ, while sympathetic neurons are responsible for releasing acetylcholine.

The autonomic nervous system serves as the relay between the CNS and the internal organs. It controls the lungs, the heart, smooth muscle, and exocrine and endocrine glands. The autonomic nervous system controls these organs largely without conscious control; it can continuously monitor the conditions of these different systems and implement changes as needed. Signaling to the target tissue usually involves two synapses: a preganglionic neuron (originating in the CNS) synapses to a neuron in a ganglion that, in turn, synapses on the target organ, as illustrated in [link]. There are two divisions of the autonomic nervous system that often have opposing effects: the sympathetic nervous system and the parasympathetic nervous system.

Sympathetic Nervous System

The **sympathetic nervous system** is responsible for the “fight or flight” response that occurs when an animal encounters a dangerous situation. One way to remember this is to think of the surprise a person

feels when encountering a snake ("snake" and "sympathetic" both begin with "s"). Examples of functions controlled by the sympathetic nervous system include an accelerated heart rate and inhibited digestion. These functions help prepare an organism's body for the physical strain required to escape a potentially dangerous situation or to fend off a predator.



Most preganglionic neurons in the sympathetic nervous system originate in the spinal cord, as illustrated in [link]. The axons of these neurons release **acetylcholine** on postganglionic neurons

within sympathetic ganglia (the sympathetic ganglia form a chain that extends alongside the spinal cord). The acetylcholine activates the postganglionic neurons. Postganglionic neurons then release **norepinephrine** onto target organs. As anyone who has ever felt a rush before a big test, speech, or athletic event can attest, the effects of the sympathetic nervous system are quite pervasive. This is both because one preganglionic neuron synapses on multiple postganglionic neurons, amplifying the effect of the original synapse, and because the adrenal gland also releases norepinephrine (and the closely related hormone epinephrine) into the bloodstream. The physiological effects of this norepinephrine release include dilating the trachea and bronchi (making it easier for the animal to breathe), increasing heart rate, and moving blood from the skin to the heart, muscles, and brain (so the animal can think and run). The strength and speed of the sympathetic response helps an organism avoid danger, and scientists have found evidence that it may also increase LTP—allowing the animal to remember the dangerous situation and avoid it in the future.

Parasympathetic Nervous System

While the sympathetic nervous system is activated in stressful situations, the **parasympathetic nervous system** allows an animal to "rest and digest." One way to remember this is to think that

during a restful situation like a picnic, the parasympathetic nervous system is in control ("picnic" and "parasympathetic" both start with "p"). Parasympathetic preganglionic neurons have cell bodies located in the brainstem and in the sacral (toward the bottom) spinal cord, as shown in [link]. The axons of the preganglionic neurons release acetylcholine on the postganglionic neurons, which are generally located very near the target organs. Most postganglionic neurons release acetylcholine onto target organs, although some release nitric oxide.

The parasympathetic nervous system resets organ function after the sympathetic nervous system is activated (the common adrenaline dump you feel after a 'fight-or-flight' event). Effects of acetylcholine release on target organs include slowing of heart rate, lowered blood pressure, and stimulation of digestion.

The human brain contains 12 cranial nerves that receive sensory input and control motor output for the head and neck. Spinal nerves contain both sensory and motor axons. The somas of sensory neurons are located in dorsal root ganglia. The somas of motor neurons are found in the ventral portion of the gray matter of the spinal cord.

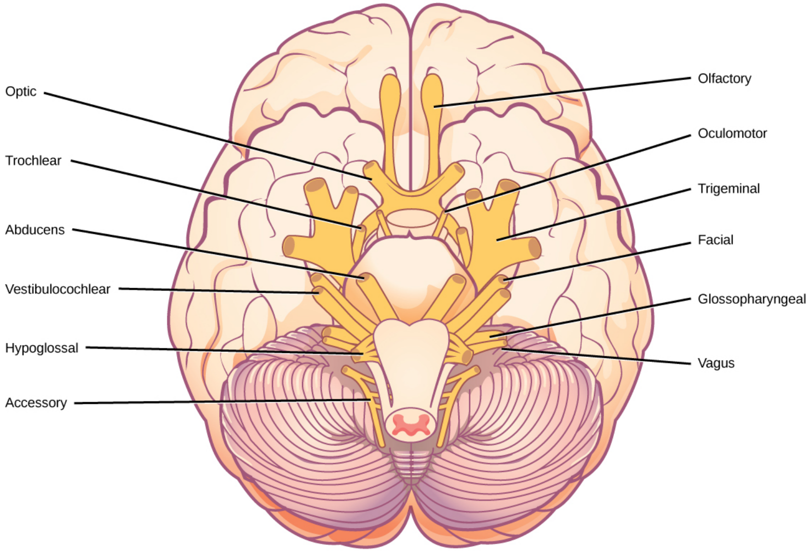
Sensory-Somatic Nervous System

The sensory-somatic nervous system is made up of

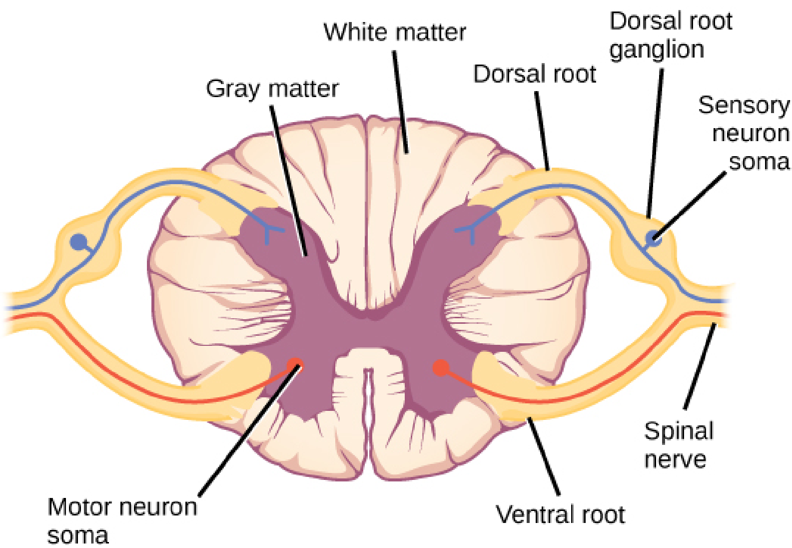
cranial and spinal nerves and contains both sensory and motor neurons. Sensory neurons transmit sensory information from the skin, skeletal muscle, and sensory organs to the CNS. Motor neurons transmit messages about desired movement from the CNS to the muscles to make them contract. Without its sensory-somatic nervous system, an animal would be unable to process any information about its environment (what it sees, feels, hears, and so on) and could not control motor movements. Unlike the autonomic nervous system, which has two synapses between the CNS and the target organ, sensory and motor neurons have only one synapse—one ending of the neuron is at the organ and the other directly contacts a CNS neuron. Acetylcholine is the main neurotransmitter released at these synapses.

Humans have 12 **cranial nerves**, nerves that emerge from or enter the skull (cranium), as opposed to the spinal nerves, which emerge from the vertebral column. Each cranial nerve is accorded a name, which are detailed in [link]. Some cranial nerves transmit only sensory information. For example, the olfactory nerve transmits information about smells from the nose to the brainstem. Other cranial nerves transmit almost solely motor information. For example, the oculomotor nerve controls the opening and closing of the eyelid and some eye movements. Other cranial nerves contain a mix of sensory and motor fibers. For example, the

glossopharyngeal nerve has a role in both taste (sensory) and swallowing (motor).



Spinal nerves transmit sensory and motor information between the spinal cord and the rest of the body. Each of the 31 spinal nerves (in humans) contains both sensory and motor axons. The sensory neuron cell bodies are grouped in structures called dorsal root ganglia and are shown in [link]. Each sensory neuron has one projection—with a sensory receptor ending in skin, muscle, or sensory organs—and another that synapses with a neuron in the dorsal spinal cord. Motor neurons have cell bodies in the ventral gray matter of the spinal cord that project to muscle through the ventral root. These neurons are usually stimulated by interneurons within the spinal cord but are sometimes directly stimulated by sensory neurons.



Section Summary

The peripheral nervous system contains both the autonomic and sensory-somatic nervous systems. The autonomic nervous system provides unconscious control over visceral functions and has two divisions: the sympathetic and parasympathetic nervous systems. The sympathetic nervous system is activated in stressful situations to prepare the animal for a “fight or flight” response. The parasympathetic nervous system is active during restful periods. The sensory-somatic nervous system is made of cranial and spinal nerves that transmit sensory information from skin and muscle to the CNS and motor commands from the CNS to the

muscles.

Visual Connection Questions

[link] Which of the following statements is false?

1. The parasympathetic pathway is responsible for relaxing the body, while the sympathetic pathway is responsible for preparing for an emergency.
2. Most preganglionic neurons in the sympathetic pathway originate in the spinal cord.
3. Slowing of the heartbeat is a parasympathetic response.
4. Parasympathetic neurons are responsible for releasing norepinephrine on the target organ, while sympathetic neurons are responsible for releasing acetylcholine.

[link] D

Review Questions

Activation of the sympathetic nervous system causes:

1. increased blood flow into the skin
2. a decreased heart rate
3. an increased heart rate
4. increased digestion

C

Where are parasympathetic preganglionic cell bodies located?

1. cerebellum
2. brainstem
3. dorsal root ganglia
4. skin

B

______ is released by motor nerve endings onto muscle.

1. Acetylcholine
2. Norepinephrine
3. Dopamine
4. Serotonin

A

Critical Thinking Questions

What are the main differences between the sympathetic and parasympathetic branches of the autonomic nervous system?

The sympathetic nervous system prepares the body for “fight or flight,” whereas the parasympathetic nervous system allows the body to “rest and digest.” Sympathetic neurons release norepinephrine onto target organs; parasympathetic neurons release acetylcholine. Sympathetic neuron cell bodies are located in sympathetic ganglia. Parasympathetic neuron cell bodies are located in the brainstem and sacral spinal cord. Activation of the sympathetic nervous system increases heart rate and blood pressure and decreases digestion and blood flow to the skin. Activation of the parasympathetic nervous system decreases heart rate and blood pressure and increases digestion and blood flow to the skin.

What are the main functions of the sensory-

somatic nervous system?

The sensory-somatic nervous system transmits sensory information from the skin, muscles, and sensory organs to the CNS. It also sends motor commands from the CNS to the muscles, causing them to contract.

Describe how the sensory-somatic nervous system reacts by reflex to a person touching something hot. How does this allow for rapid responses in potentially dangerous situations?

A person's skin comes into contact with a hot object, and the high temperature is recognized by the thermoreceptors of a sensory neuron. The signal is relayed to the spinal cord, and sent to a motor neuron. The motor neuron relays the signal to its axon, and produces acetylcholine to contract the muscle that will pull the person away from the hot object. By connecting the sensory and motor neurons in the spinal cord (instead of integrating the signal in the brain) the body can respond faster.

Scientists have suggested that the autonomic nervous system is not well-adapted to modern human life. How is the sympathetic nervous

system an ineffective response to the everyday challenges faced by modern humans?

Many events in modern human life are not physical dangers; instead they are events we think of as “stress.” Finding the money to pay your student loans or being nervous before a test still activate the sympathetic nervous system, but these situations do not require the fight-or-flight response to survive.

Glossary

acetylcholine
: neurotransmitter released by neurons in the central nervous system and peripheral nervous system

autonomic nervous system
: part of the peripheral nervous system that controls bodily functions

cranial nerve
: sensory and/or motor nerve that emanates from the brain

norepinephrine
: neurotransmitter and hormone released by activation of the sympathetic nervous system

parasympathetic nervous system
: division of autonomic nervous system that regulates visceral functions during rest and digestion

sensory-somatic nervous system
: system of sensory and motor nerves

spinal nerve
: nerve projecting between skin or muscle and spinal cord

sympathetic nervous system
: division of autonomic nervous system activated during stressful “fight or flight” situations

Nervous System Disorders
By the end of this section, you will be able to do the following:

- Describe the symptoms, potential causes, and treatment of several examples of nervous system disorders

A nervous system that functions correctly is a fantastically complex, well-oiled machine—synapses fire appropriately, muscles move when needed, memories are formed and stored, and emotions are well regulated. Unfortunately, each year millions of people in the United States deal with some sort of nervous system disorder. While scientists have discovered potential causes of many of these diseases, and viable treatments for some, ongoing research seeks to find ways to better prevent and treat all of these disorders.
Compared to a normal brain (left), the brain from a patient with Alzheimer's disease (right) shows a dramatic neurodegeneration, particularly within the ventricles and hippocampus. (credit: modification of work by "Garrando"/Wikimedia Commons based on original images by ADEAR: "Alzheimer's Disease Education and Referral Center, a service of the National Institute on Aging") Parkinson's patients often have a characteristic hunched walk.

Neurodegenerative Disorders

Neurodegenerative disorders are illnesses characterized by a loss of nervous system functioning that are usually caused by neuronal death. These diseases generally worsen over time as more and more neurons die. The symptoms of a particular neurodegenerative disease are related to where in the nervous system the death of neurons occurs. Spinocerebellar ataxia, for example, leads to neuronal death in the cerebellum. The death of these neurons causes problems in balance and walking. Neurodegenerative disorders include Huntington's disease, amyotrophic lateral sclerosis, Alzheimer's disease and other types of dementia disorders, and Parkinson's disease. Here, Alzheimer's and Parkinson's disease will be discussed in more depth.

Alzheimer's Disease

Alzheimer's disease is the most common cause of dementia in the elderly. In 2012, an estimated 5.4 million Americans suffered from Alzheimer's disease, and payments for their care are estimated at $200 billion. Roughly one in every eight people age 65 or older has the disease. Due to the aging of the baby-boomer generation, there are projected to be as many as 13 million Alzheimer's patients in the United States in the year 2050.

Symptoms of Alzheimer's disease include disruptive memory loss, confusion about time or place,

difficulty planning or executing tasks, poor judgment, and personality changes. Problems smelling certain scents can also be indicative of Alzheimer's disease and may serve as an early warning sign. Many of these symptoms are also common in people who are aging normally, so it is the severity and longevity of the symptoms that determine whether a person is suffering from Alzheimer's.

Alzheimer's disease was named for Alois Alzheimer, a German psychiatrist who published a report in 1911 about a woman who showed severe dementia symptoms. Along with his colleagues, he examined the woman's brain following her death and reported the presence of abnormal clumps, which are now called amyloid plaques, along with tangled brain fibers called neurofibrillary tangles. Amyloid plaques, neurofibrillary tangles, and an overall shrinking of brain volume are commonly seen in the brains of Alzheimer's patients. Loss of neurons in the hippocampus is especially severe in advanced Alzheimer's patients. [link] compares a normal brain to the brain of an Alzheimer's patient. Many research groups are examining the causes of these hallmarks of the disease.

One form of the disease is usually caused by mutations in one of three known genes. This rare form of early onset Alzheimer's disease affects fewer than five percent of patients with the disease and

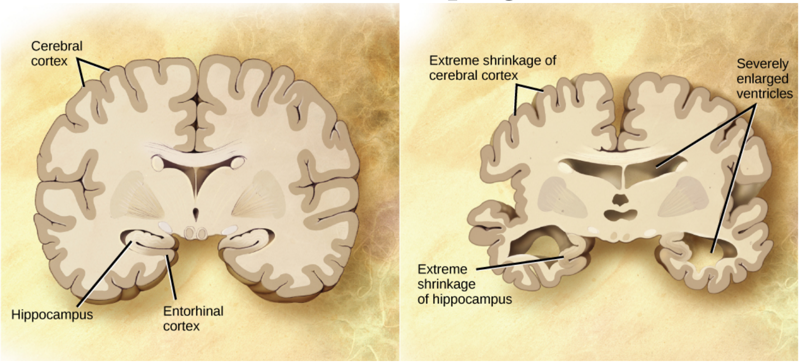
causes dementia beginning between the ages of 30 and 60. The more prevalent, late-onset form of the disease likely also has a genetic component. One particular gene, apolipoprotein E (APOE) has a variant (E4) that increases a carrier's likelihood of getting the disease. Many other genes have been identified that might be involved in the pathology.

Link to Learning

Visit this website for video links discussing genetics and Alzheimer's disease.

Unfortunately, there is no cure for Alzheimer's disease. Current treatments focus on managing the symptoms of the disease. Because decrease in the activity of cholinergic neurons (neurons that use the neurotransmitter acetylcholine) is common in Alzheimer's disease, several drugs used to treat the disease work by increasing acetylcholine neurotransmission, often by inhibiting the enzyme that breaks down acetylcholine in the synaptic cleft. Other clinical interventions focus on behavioral therapies like psychotherapy, sensory therapy, and cognitive exercises. Since Alzheimer's disease appears to hijack the normal aging process, research into prevention is prevalent. Smoking, obesity, and cardiovascular problems may be risk factors for the

disease, so treatments for those may also help to prevent Alzheimer's disease. Some studies have shown that people who remain intellectually active by playing games, reading, playing musical instruments, and being socially active in later life have a reduced risk of developing the disease.



Parkinson's Disease

Like Alzheimer's disease, **Parkinson's disease** is a neurodegenerative disease. It was first characterized by James Parkinson in 1817. Each year, 50,000-60,000 people in the United States are diagnosed with the disease. Parkinson's disease causes the loss of dopamine neurons in the substantia nigra, a midbrain structure that regulates movement. Loss of these neurons causes many symptoms including tremor (shaking of fingers or a limb), slowed movement, speech changes, balance and posture problems, and rigid muscles. The combination of these symptoms often causes a characteristic slow hunched shuffling walk, illustrated in [link]. Patients with Parkinson's

disease can also exhibit psychological symptoms, such as dementia or emotional problems.

Although some patients have a form of the disease known to be caused by a single mutation, for most patients the exact causes of Parkinson's disease remain unknown: the disease likely results from a combination of genetic and environmental factors (similar to Alzheimer's disease). Post-mortem analysis of brains from Parkinson's patients shows the presence of Lewy bodies—abnormal protein clumps—in dopaminergic neurons. The prevalence of these Lewy bodies often correlates with the severity of the disease.

There is no cure for Parkinson's disease, and treatment is focused on easing symptoms. One of the most commonly prescribed drugs for Parkinson's is L-DOPA, which is a chemical that is converted into dopamine by neurons in the brain. This conversion increases the overall level of dopamine neurotransmission and can help compensate for the loss of dopaminergic neurons in the substantia nigra. Other drugs work by inhibiting the enzyme that breaks down dopamine.

Many people with ADHD have one or more other neurological disorders. (credit “chart design and illustration”: modification of work by Leigh Coriale; credit “data”: Drs. Biederman and Faraone, Massachusetts General Hospital).

Neurodevelopmental Disorders

Neurodevelopmental disorders occur when the development of the nervous system is disturbed. There are several different classes of neurodevelopmental disorders. Some, like Down

Syndrome, cause intellectual deficits. Others specifically affect communication, learning, or the motor system. Some disorders like autism spectrum disorder and attention deficit/hyperactivity disorder have complex symptoms.

Autism

Autism spectrum disorder (ASD) is a neurodevelopmental disorder. Its severity differs from person to person. Estimates for the prevalence of the disorder have changed rapidly in the past few decades. Current estimates suggest that one in 88 children will develop the disorder. ASD is four times more prevalent in males than females.

Link to Learning

This video discusses possible reasons why there has been a recent increase in the number of people diagnosed with autism.

A characteristic symptom of ASD is impaired social skills. Children with autism may have difficulty making and maintaining eye contact and reading social cues. They also may have problems feeling empathy for others. Other symptoms of ASD include repetitive motor behaviors (such as rocking back

and forth), preoccupation with specific subjects, strict adherence to certain rituals, and unusual language use. Up to 30 percent of patients with ASD develop epilepsy, and patients with some forms of the disorder (like Fragile X) also have intellectual disability. Because it is a spectrum disorder, other ASD patients are very functional and have good-to-excellent language skills. Many of these patients do not feel that they suffer from a disorder and instead think that their brains just process information differently.

Except for some well-characterized, clearly genetic forms of autism (like Fragile X and Rett's Syndrome), the causes of ASD are largely unknown. Variants of several genes correlate with the presence of ASD, but for any given patient, many different mutations in different genes may be required for the disease to develop. At a general level, ASD is thought to be a disease of "incorrect" wiring. Accordingly, brains of some ASD patients lack the same level of synaptic pruning that occurs in non-affected people. In the 1990s, a research paper linked autism to a common vaccine given to children. This paper was retracted when it was discovered that the author falsified data, and follow-up studies showed no connection between vaccines and autism.

Treatment for autism usually combines behavioral therapies and interventions, along with medications

to treat other disorders common to people with autism (depression, anxiety, obsessive compulsive disorder). Although early interventions can help mitigate the effects of the disease, there is currently no cure for ASD.

Attention Deficit Hyperactivity Disorder (ADHD)

Approximately three to five percent of children and adults are affected by **attention deficit/ hyperactivity disorder (ADHD)**. Like ASD, ADHD is more prevalent in males than females. Symptoms of the disorder include inattention (lack of focus), executive functioning difficulties, impulsivity, and hyperactivity beyond what is characteristic of the normal developmental stage. Some patients do not have the hyperactive component of symptoms and are diagnosed with a subtype of ADHD: attention deficit disorder (ADD). Many people with ADHD also show comorbitity, in that they develop secondary disorders in addition to ADHD. Examples include depression or obsessive compulsive disorder (OCD). [link] provides some statistics concerning comorbidity with ADHD.

The cause of ADHD is unknown, although research points to a delay and dysfunction in the development of the prefrontal cortex and disturbances in neurotransmission. According to studies of twins, the disorder has a strong genetic component. There are several candidate genes that

may contribute to the disorder, but no definitive links have been discovered. Environmental factors, including exposure to certain pesticides, may also contribute to the development of ADHD in some patients. Treatment for ADHD often involves behavioral therapies and the prescription of stimulant medications, which paradoxically cause a calming effect in these patients.

Career Connection

Neurologist

Neurologists are physicians who specialize in disorders of the nervous system. They diagnose and treat disorders such as epilepsy, stroke, dementia, nervous system injuries, Parkinson's disease, sleep disorders, and multiple sclerosis. Neurologists are medical doctors who have attended college, medical school, and completed three to four years of neurology residency.

When examining a new patient, a neurologist takes a full medical history and performs a complete physical exam. The physical exam contains specific tasks that are used to determine what areas of the

brain, spinal cord, or peripheral nervous system may be damaged. For example, to check whether the hypoglossal nerve is functioning correctly, the neurologist will ask the patient to move his or her tongue in different ways. If the patient does not have full control over tongue movements, then the hypoglossal nerve may be damaged or there may be a lesion in the brainstem where the cell bodies of these neurons reside (or there could be damage to the tongue muscle itself).

Neurologists have other tools besides a physical exam they can use to diagnose particular problems in the nervous system. If the patient has had a seizure, for example, the neurologist can use electroencephalography (EEG), which involves taping electrodes to the scalp to record brain activity, to try to determine which brain regions are involved in the seizure. In suspected stroke patients, a neurologist can use a computerized tomography (CT) scan, which is a type of X-ray, to look for bleeding in the brain or a possible brain tumor. To treat patients with neurological problems, neurologists can prescribe medications or refer the patient to a neurosurgeon for surgery.

Link to Learning

This website allows you to see the different tests a neurologist might use to see what regions of the nervous system may be damaged in a patient.

Mental Illnesses

Mental illnesses are nervous system disorders that result in problems with thinking, mood, or relating with other people. These disorders are severe enough to affect a person's quality of life and often make it difficult for people to perform the routine tasks of daily living. Debilitating mental disorders plague approximately 12.5 million Americans (about 1 in 17 people) at an annual cost of more than $300 billion. There are several types of mental disorders including schizophrenia, major depression, bipolar disorder, anxiety disorders and phobias, post-traumatic stress disorders, and obsessive-compulsive disorder (OCD), among others. The American Psychiatric Association publishes the Diagnostic and Statistical Manual of Mental Disorders (or DSM), which describes the symptoms required for a patient to be diagnosed with a particular mental disorder. Each newly released version of the DSM contains different symptoms and classifications as scientists learn more about these disorders, their causes, and how they relate to each other. A more detailed discussion of two mental illnesses—schizophrenia and major depression—is given below.

Schizophrenia

Schizophrenia is a serious and often debilitating mental illness affecting one percent of people in the

United States. Symptoms of the disease include the inability to differentiate between reality and imagination, inappropriate and unregulated emotional responses, difficulty thinking, and problems with social situations. People with schizophrenia can suffer from hallucinations and hear voices; they may also suffer from delusions. Patients also have so-called "negative" symptoms like a flattened emotional state, loss of pleasure, and loss of basic drives. Many schizophrenic patients are diagnosed in their late adolescence or early 20s. The development of schizophrenia is thought to involve malfunctioning dopaminergic neurons and may also involve problems with glutamate signaling. Treatment for the disease usually requires antipsychotic medications that work by blocking dopamine receptors and decreasing dopamine neurotransmission in the brain. This decrease in dopamine can cause Parkinson's disease-like symptoms in some patients. While some classes of antipsychotics can be quite effective at treating the disease, they are not a cure, and most patients must remain medicated for the rest of their lives.

Depression

Major depression affects approximately 6.7 percent of the adults in the United States each year and is one of the most common mental disorders. To be diagnosed with major depressive disorder, a person must have experienced a severely depressed mood

lasting longer than two weeks along with other symptoms including a loss of enjoyment in activities that were previously enjoyed, changes in appetite and sleep schedules, difficulty concentrating, feelings of worthlessness, and suicidal thoughts. The exact causes of major depression are unknown and likely include both genetic and environmental risk factors. Some research supports the “classic monoamine hypothesis,” which suggests that depression is caused by a decrease in norepinephrine and serotonin neurotransmission. One argument against this hypothesis is the fact that some antidepressant medications cause an increase in norepinephrine and serotonin release within a few hours of beginning treatment—but clinical results of these medications are not seen until weeks later. This has led to alternative hypotheses: for example, dopamine may also be decreased in depressed patients, or it may actually be an increase in norepinephrine and serotonin that causes the disease, and antidepressants force a feedback loop that decreases this release. Treatments for depression include psychotherapy, electroconvulsive therapy, deep-brain stimulation, and prescription medications. There are several classes of antidepressant medications that work through different mechanisms. For example, monoamine oxidase inhibitors (MAO inhibitors) block the enzyme that degrades many neurotransmitters (including dopamine, serotonin, norepinephrine), resulting in increased neurotransmitter in the

synaptic cleft. Selective serotonin reuptake inhibitors (SSRIs) block the reuptake of serotonin into the presynaptic neuron. This blockage results in an increase in serotonin in the synaptic cleft. Other types of drugs such as norepinephrine-dopamine reuptake inhibitors and norepinephrine-serotonin reuptake inhibitors are also used to treat depression.

Other Neurological Disorders

There are several other neurological disorders that cannot be easily placed in the above categories. These include chronic pain conditions, cancers of the nervous system, epilepsy disorders, and stroke. Epilepsy and stroke are discussed below.

Epilepsy

Estimates suggest that up to three percent of people in the United States will be diagnosed with **epilepsy** in their lifetime. While there are several different types of epilepsy, all are characterized by recurrent seizures. Epilepsy itself can be a symptom of a brain injury, disease, or other illness. For example, people who have intellectual disability or ASD can experience seizures, presumably because the developmental wiring malfunctions that caused their disorders also put them at risk for epilepsy. For many patients, however, the cause of their epilepsy

is never identified and is likely to be a combination of genetic and environmental factors. Often, seizures can be controlled with anticonvulsant medications. However, for very severe cases, patients may undergo brain surgery to remove the brain area where seizures originate.

Stroke

A stroke results when blood fails to reach a portion of the brain for a long enough time to cause damage. Without the oxygen supplied by blood flow, neurons in this brain region die. This neuronal death can cause many different symptoms—depending on the brain area affected— including headache, muscle weakness or paralysis, speech disturbances, sensory problems, memory loss, and confusion. Stroke is often caused by blood clots and can also be caused by the bursting of a weak blood vessel. Strokes are extremely common and are the third most common cause of death in the United States. On average one person experiences a stroke every 40 seconds in the United States. Approximately 75 percent of strokes occur in people older than 65. Risk factors for stroke include high blood pressure, diabetes, high cholesterol, and a family history of stroke. Smoking doubles the risk of stroke. Because a stroke is a medical emergency, patients with symptoms of a stroke should immediately go to the emergency room, where they can receive drugs that will dissolve any clot that

may have formed. These drugs will not work if the stroke was caused by a burst blood vessel or if the stroke occurred more than three hours before arriving at the hospital. Treatment following a stroke can include blood pressure medication (to prevent future strokes) and (sometimes intense) physical therapy.

Section Summary

Some general themes emerge from the sampling of nervous system disorders presented above. The causes for most disorders are not fully understood—at least not for all patients—and likely involve a combination of nature (genetic mutations that become risk factors) and nurture (emotional trauma, stress, hazardous chemical exposure). Because the causes have yet to be fully determined, treatment options are often lacking and only address symptoms.

Review Questions

Parkinson's disease is a caused by the degeneration of neurons that release _____.

1. serotonin

2. dopamine
3. glutamate
4. norepinephrine

B

______ medications are often used to treat patients with ADHD.

1. Tranquilizer
2. Antibiotic
3. Stimulant
4. Anti-seizure

C

Strokes are often caused by ______.

1. neurodegeneration
2. blood clots or burst blood vessels
3. seizures
4. viruses

B

Why is it difficult to identify the cause of many

nervous system disorders?

1. The genes associated with the diseases are not known.
2. There are no obvious defects in brain structure.
3. The onset and display of symptoms varies between patients.
4. all of the above

D

Why do many patients with neurodevelopmental disorders develop secondary disorders?

1. Their genes predispose them to schizophrenia.
2. Stimulant medications cause new behavioral disorders.
3. Behavioral therapies only improve neurodevelopmental disorders.
4. Dysfunction in the brain can affect many aspects of the body.

D

Critical Thinking Questions

What are the main symptoms of Alzheimer's disease?

Symptoms of Alzheimer's disease include disruptive memory loss, confusion about time or place, difficulties planning or executing tasks, poor judgment, and personality changes.

What are possible treatments for patients with major depression?

Possible treatments for patients with major depression include psychotherapy and prescription medications. MAO inhibitor drugs inhibit the breakdown of certain neurotransmitters (including dopamine, serotonin, norepinephrine) in the synaptic cleft. SSRI medications inhibit the reuptake of serotonin into the presynaptic neuron.

Glossary

Alzheimer's disease
: neurodegenerative disorder characterized by

problems with memory and thinking

attention deficit hyperactivity disorder (ADHD)
: neurodevelopmental disorder characterized by difficulty maintaining attention and controlling impulses

autism spectrum disorder (ASD)
: neurodevelopmental disorder characterized by impaired social interaction and communication abilities

epilepsy
: neurological disorder characterized by recurrent seizures

major depression
: mental illness characterized by prolonged periods of sadness

neurodegenerative disorder
: nervous system disorder characterized by the progressive loss of neurological functioning, usually caused by neuron death

Parkinson's disease
: neurodegenerative disorder that affects the control of movement

schizophrenia
: mental disorder characterized by the inability to accurately perceive reality; patients often

have difficulty thinking clearly and can suffer from delusions

Introduction

class="introduction" Improvements in the design of prostheses have allowed for a wider range of activities in recipients. (credit: modification of work by Stuart Grout)

The muscular and skeletal systems provide support to the body and allow for a wide range of movement. The bones of the skeletal system protect the body’s internal organs and support the weight of the body. The muscles of the muscular system contract and pull on the bones, allowing for movements as diverse as standing, walking, running, and grasping items.

Injury or disease affecting the musculoskeletal system can be very debilitating. In humans, the most common musculoskeletal diseases worldwide are caused by malnutrition. Ailments that affect the

joints are also widespread, such as arthritis, which can make movement difficult and—in advanced cases—completely impair mobility. In severe cases in which the joint has suffered extensive damage, joint replacement surgery may be needed.

Progress in the science of prosthesis design has resulted in the development of artificial joints, with joint replacement surgery in the hips and knees being the most common. Replacement joints for shoulders, elbows, and fingers are also available. Even with this progress, there is still room for improvement in the design of prostheses. The state-of-the-art prostheses have limited durability and therefore wear out quickly, particularly in young or active individuals. Current research is focused on the use of new materials, such as carbon fiber, that may make prostheses more durable.

Types of Skeletal Systems
By the end of this section, you will be able to do the following:

- Discuss the different types of skeletal systems
- Explain the role of the human skeletal system
- Compare and contrast different skeletal systems

A skeletal system is necessary to support the body, protect internal organs, and allow for the movement of an organism. There are three different skeleton designs that fulfill these functions: hydrostatic skeleton, exoskeleton, and endoskeleton.
The skeleton of the red-knobbed sea star (*Protoreaster linckii*) is an example of a hydrostatic skeleton. (credit: “Amada44”/Wikimedia Commons)

Hydrostatic Skeleton

A **hydrostatic skeleton** is a skeleton formed by a fluid-filled compartment within the body, called the coelom. The organs of the coelom are supported by the aqueous fluid, which also resists external compression. This compartment is under hydrostatic pressure because of the fluid and supports the other organs of the organism. This type of skeletal system is found in soft-bodied animals such as sea anemones, earthworms, Cnidaria, and other invertebrates ([link]).

Movement in a hydrostatic skeleton is provided by muscles that surround the coelom. The muscles in a hydrostatic skeleton contract to change the shape of the coelom; the pressure of the fluid in the coelom produces movement. For example, earthworms move by waves of muscular contractions of the skeletal muscle of the body wall hydrostatic skeleton, called peristalsis, which alternately shorten and lengthen the body. Lengthening the body extends the anterior end of the organism. Most organisms have a mechanism to fix themselves in the substrate. Shortening the muscles then draws the posterior portion of the body forward. Although a hydrostatic skeleton is well-suited to invertebrate organisms such as earthworms and some aquatic organisms, it is not an efficient skeleton for terrestrial animals.

Muscles attached to the exoskeleton of the Halloween crab (*Gecarcinus quadratus*) allow it to

move.

Exoskeleton

An **exoskeleton** is an external skeleton that consists of a hard encasement on the surface of an organism. For example, the shells of crabs and insects are exoskeletons ([link]). This skeleton type provides defence against predators, supports the body, and allows for movement through the contraction of attached muscles. As with vertebrates, muscles must cross a joint inside the exoskeleton. Shortening of the muscle changes the relationship of the two segments of the exoskeleton. Arthropods such as crabs and lobsters have exoskeletons that consist of 30–50 percent chitin, a polysaccharide derivative of glucose that is a strong but flexible material. Chitin is secreted by the epidermal cells. The exoskeleton is further strengthened by the addition of calcium carbonate in organisms such as the lobster. Because the exoskeleton is acellular, arthropods must periodically shed their exoskeletons because the exoskeleton does not grow as the organism grows.

The skeletons of humans and horses are examples of endoskeletons. (credit: Ross Murphy)

Endoskeleton

An **endoskeleton** is a skeleton that consists of hard, mineralized structures located within the soft tissue of organisms. An example of a primitive endoskeletal structure is the spicules of sponges. The bones of vertebrates are composed of tissues, whereas sponges have no true tissues ([link]). Endoskeletons provide support for the body, protect internal organs, and allow for movement through contraction of muscles attached to the skeleton.

The human skeleton is an endoskeleton that consists of 206 bones in the adult. It has five main functions: providing support to the body, storing minerals and lipids, producing blood cells, protecting internal organs, and allowing for movement. The skeletal system in vertebrates is divided into the axial skeleton (which consists of the skull, vertebral column, and rib cage), and the appendicular

skeleton (which consists of the shoulders, limb bones, the pectoral girdle, and the pelvic girdle).

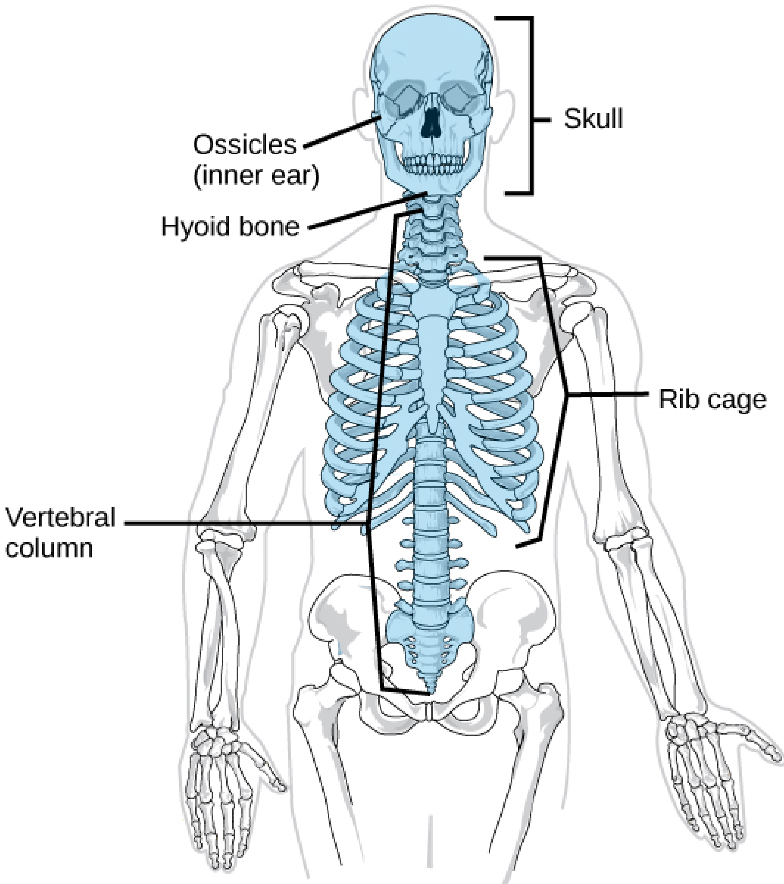
Link to Learning
Visit the interactive body site to build a virtual skeleton: select "skeleton" and click through the activity to place each bone.

The axial skeleton consists of the bones of the skull, ossicles of the middle ear, hyoid bone, vertebral column, and rib cage. (credit: modification of work by Mariana Ruiz Villareal) The bones of the skull support the structures of the face and protect the brain. (credit: modification of work by Mariana Ruiz Villareal) The cranial bones, including the frontal, parietal, and sphenoid bones, cover the top of the head. The facial bones of the skull form the face and provide cavities for the eyes, nose, and mouth. (a) The vertebral column consists of seven cervical vertebrae (C1–7) twelve thoracic vertebrae (Th1–12), five lumbar vertebrae (L1–5), the os sacrum, and the coccyx. (b) Spinal curves increase the strength and flexibility of the spine. (credit a: modification of work by Uwe Gille based on original work by Gray's Anatomy; credit b: modification of work by NCI, NIH) The thoracic cage, or rib cage, protects the heart and the lungs. (credit:

modification of work by NCI, NIH)

Human Axial Skeleton

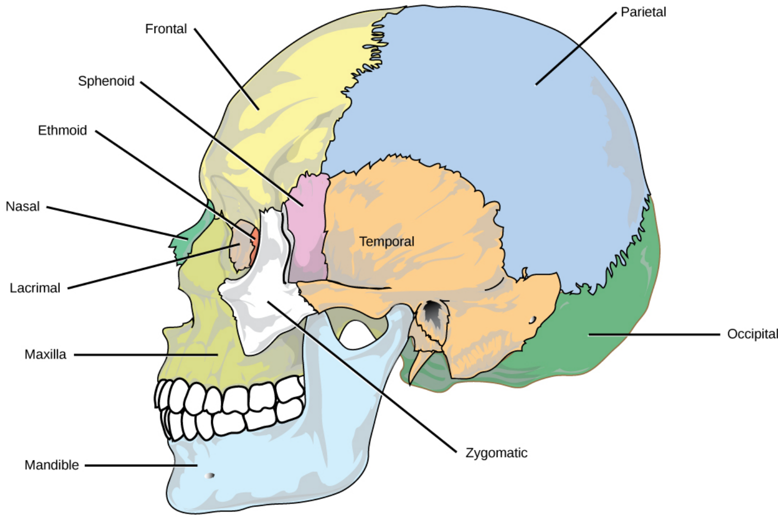
The **axial skeleton** forms the central axis of the body and includes the bones of the skull, ossicles of the middle ear, hyoid bone of the throat, vertebral column, and the thoracic cage (ribcage) ([link]). The function of the axial skeleton is to provide support and protection for the brain, the spinal cord, and the organs in the ventral body cavity. It provides a surface for the attachment of muscles that move the head, neck, and trunk, performs respiratory movements, and stabilizes parts of the appendicular skeleton.



The Skull

The bones of the **skull** support the structures of the face and protect the brain. The skull consists of 22 bones, which are divided into two categories: cranial bones and facial bones. The **cranial bones** are eight bones that form the cranial cavity, which encloses the brain and serves as an attachment site for the muscles of the head and neck. The eight cranial bones are the frontal bone, two parietal

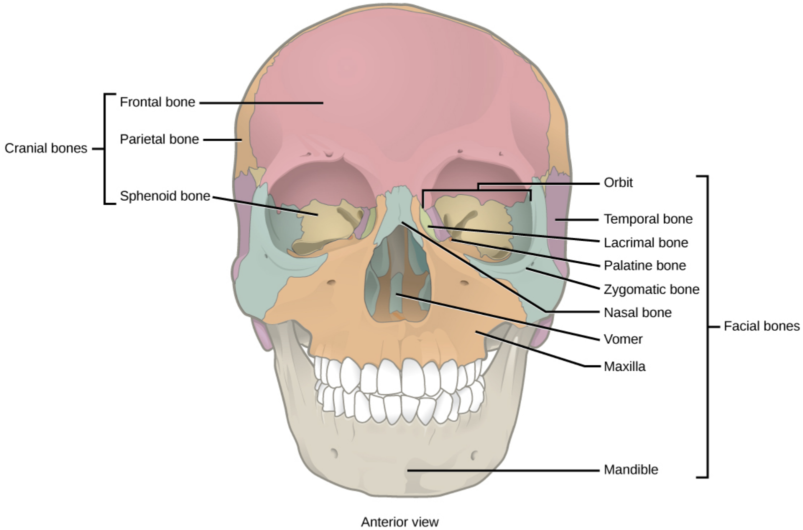
bones, two temporal bones, occipital bone, sphenoid bone, and the ethmoid bone. Although the bones developed separately in the embryo and fetus, in the adult, they are tightly fused with connective tissue and adjoining bones do not move ([link]).



The **auditory ossicles** of the middle ear transmit sounds from the air as vibrations to the fluid-filled cochlea. The auditory ossicles consist of three bones each: the malleus, incus, and stapes. These are the smallest bones in the body and are unique to mammals.

Fourteen **facial bones** form the face, provide cavities for the sense organs (eyes, mouth, and nose), protect the entrances to the digestive and respiratory tracts, and serve as attachment points for facial muscles. The 14 facial bones are the nasal bones, the maxillary bones, zygomatic bones, palatine, vomer, lacrimal bones, the inferior nasal

conchae, and the mandible. All of these bones occur in pairs except for the mandible and the vomer ([link]).

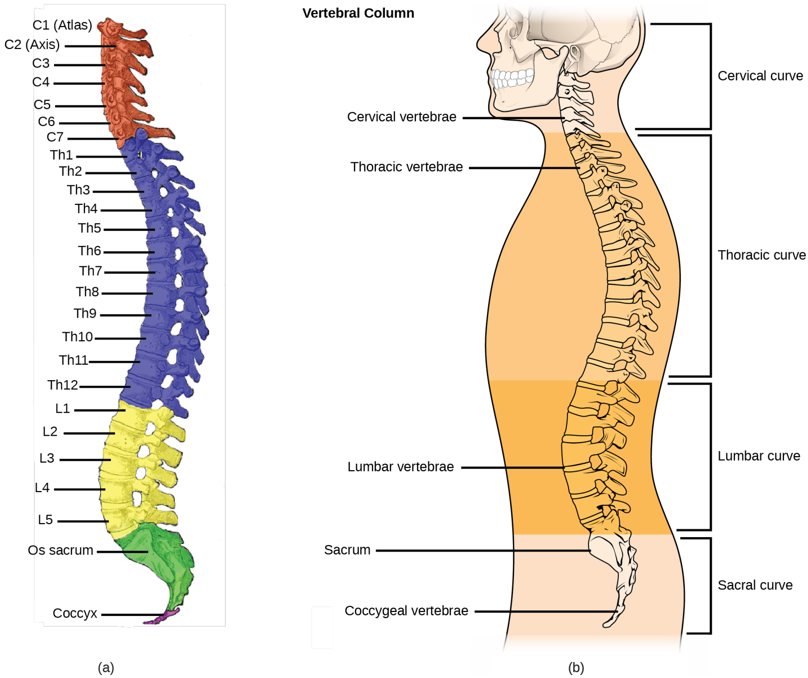


Although it is not found in the skull, the hyoid bone is considered a component of the axial skeleton. The **hyoid bone** lies below the mandible in the front of the neck. It acts as a movable base for the tongue and is connected to muscles of the jaw, larynx, and tongue. The mandible articulates with the base of the skull. The mandible controls the opening to the airway and gut. In animals with teeth, the mandible brings the surfaces of the teeth in contact with the maxillary teeth.

The Vertebral Column

The **vertebral column**, or spinal column, surrounds and protects the spinal cord, supports the head, and

acts as an attachment point for the ribs and muscles of the back and neck. The adult vertebral column comprises 26 bones: the 24 vertebrae, the sacrum, and the coccyx bones. In the adult, the sacrum is typically composed of five vertebrae that fuse into one. The coccyx is typically 3–4 vertebrae that fuse into one. Around the age of 70, the sacrum and the coccyx may fuse together. We begin life with approximately 33 vertebrae, but as we grow, several vertebrae fuse together. The adult vertebrae are further divided into the 7 cervical vertebrae, 12 thoracic vertebrae, and 5 lumbar vertebrae ([link]).



Each vertebral body has a large hole in the center through which the nerves of the spinal cord pass. There is also a notch on each side through which

the spinal nerves, which serve the body at that level, can exit from the spinal cord. The vertebral column is approximately 71 cm (28 inches) in adult male humans and is curved, which can be seen from a side view. The names of the spinal curves correspond to the region of the spine in which they occur. The thoracic and sacral curves are concave (curve inwards relative to the front of the body) and the cervical and lumbar curves are convex (curve outwards relative to the front of the body). The arched curvature of the vertebral column increases its strength and flexibility, allowing it to absorb shocks like a spring ([link]).

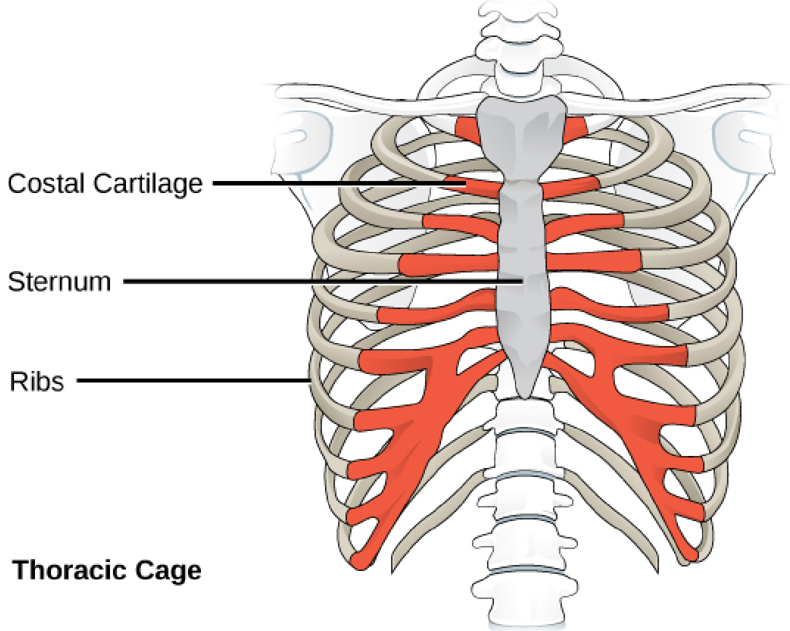
Intervertebral discs composed of fibrous cartilage lie between adjacent vertebral bodies from the second cervical vertebra to the sacrum. Each disc is part of a joint that allows for some movement of the spine and acts as a cushion to absorb shocks from movements such as walking and running. Intervertebral discs also act as ligaments to bind vertebrae together. The inner part of discs, the nucleus pulposus, hardens as people age and becomes less elastic. This loss of elasticity diminishes its ability to absorb shocks.

The Thoracic Cage

The **thoracic cage**, also known as the ribcage, is the skeleton of the chest, and consists of the ribs, sternum, thoracic vertebrae, and costal cartilages

([link]). The thoracic cage encloses and protects the organs of the thoracic cavity, including the heart and lungs. It also provides support for the shoulder girdles and upper limbs, and serves as the attachment point for the diaphragm, muscles of the back, chest, neck, and shoulders. Changes in the volume of the thorax enable breathing.

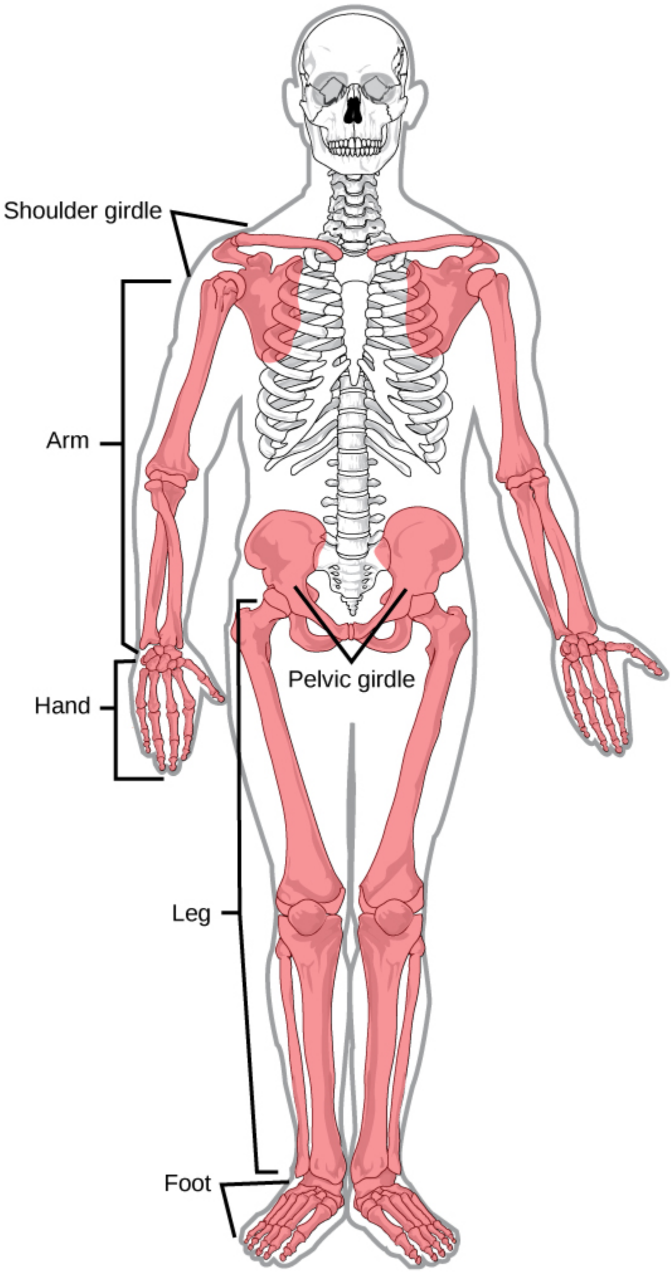
The **sternum**, or breastbone, is a long, flat bone located at the anterior of the chest. It is formed from three bones that fuse in the adult. The **ribs** are 12 pairs of long, curved bones that attach to the thoracic vertebrae and curve toward the front of the body, forming the ribcage. Costal cartilages connect the anterior ends of the ribs to the sternum, with the exception of rib pairs 11 and 12, which are free-floating ribs.



The appendicular skeleton is composed of the bones of the pectoral limbs (arm, forearm, hand), the pelvic limbs (thigh, leg, foot), the pectoral girdle, and the pelvic girdle. (credit: modification of work by Mariana Ruiz Villareal) (a) The pectoral girdle in primates consists of the clavicles and scapulae. (b) The posterior view reveals the spine of the scapula to which muscle attaches. The upper limb consists of the humerus of the upper arm, the radius and ulna of the forearm, eight bones of the carpus, five bones of the metacarpus, and 14 bones of the phalanges. To adapt to reproductive fitness, the (a) female pelvis is lighter, wider, shallower, and has a broader angle between the pubic bones than (b) the male pelvis. The lower limb consists of the thigh (femur), kneecap (patella), leg (tibia and fibula), ankle (tarsals), and foot (metatarsals and phalanges) bones. This drawing shows the bones of the human foot and ankle, including the metatarsals and the phalanges.

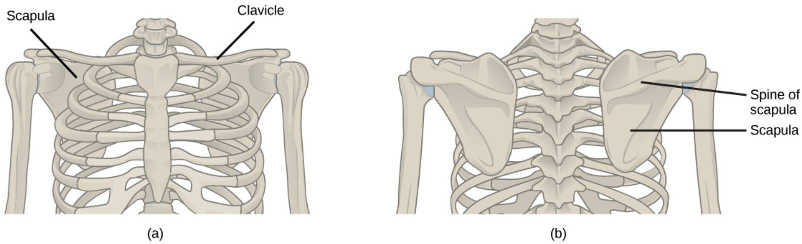
Human Appendicular Skeleton

The **appendicular skeleton** is composed of the bones of the upper limbs (which function to grasp and manipulate objects) and the lower limbs (which permit locomotion). It also includes the pectoral girdle, or shoulder girdle, that attaches the upper limbs to the body, and the pelvic girdle that attaches the lower limbs to the body ([link]).



The Pectoral Girdle

The **pectoral girdle** bones provide the points of attachment of the upper limbs to the axial skeleton. The human pectoral girdle consists of the clavicle (or collarbone) in the anterior, and the scapula (or shoulder blades) in the posterior ([link]).

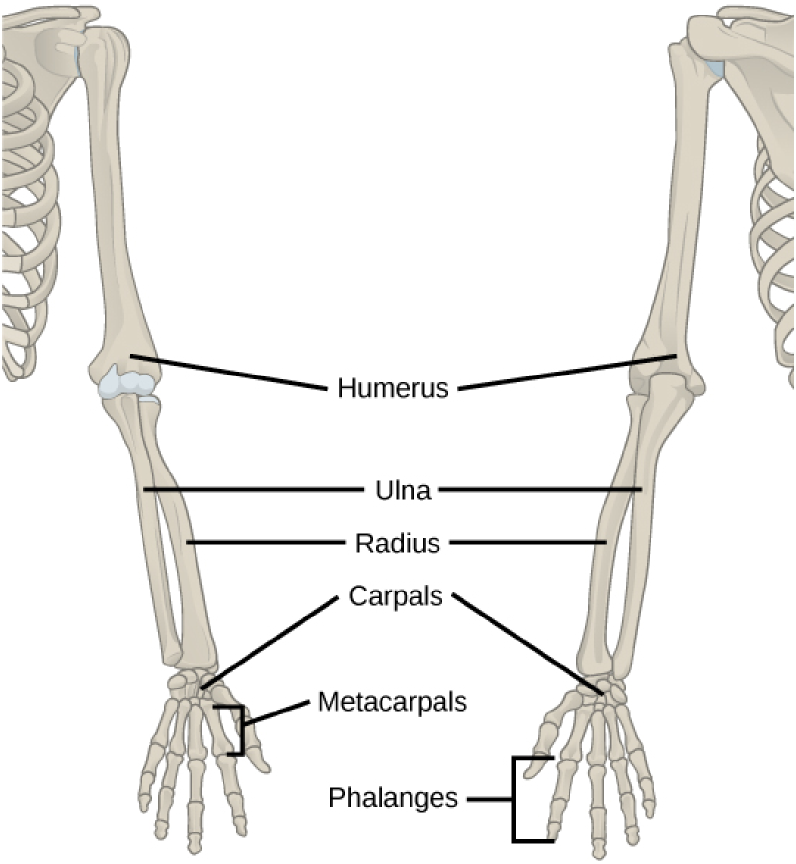


The **clavicles** are S-shaped bones that position the arms on the body. The clavicles lie horizontally across the front of the thorax (chest) just above the first rib. These bones are fairly fragile and are susceptible to fractures. For example, a fall with the arms outstretched causes the force to be transmitted to the clavicles, which can break if the force is excessive. The clavicle articulates with the sternum and the scapula.

The **scapulae** are flat, triangular bones that are located at the back of the pectoral girdle. They support the muscles crossing the shoulder joint. A ridge, called the spine, runs across the back of the scapula and can easily be felt through the skin ([link]). The spine of the scapula is a good example of a bony protrusion that facilitates a broad area of attachment for muscles to bone.

The Upper Limb

The upper limb contains 30 bones in three regions: the arm (shoulder to elbow), the forearm (ulna and radius), and the wrist and hand ([link]).



An **articulation** is any place at which two bones are joined. The **humerus** is the largest and longest bone of the upper limb and the only bone of the arm. It articulates with the scapula at the shoulder and with the forearm at the elbow. The **forearm** extends from the elbow to the wrist and consists of two bones: the ulna and the radius. The **radius** is located along the lateral (thumb) side of the forearm and articulates

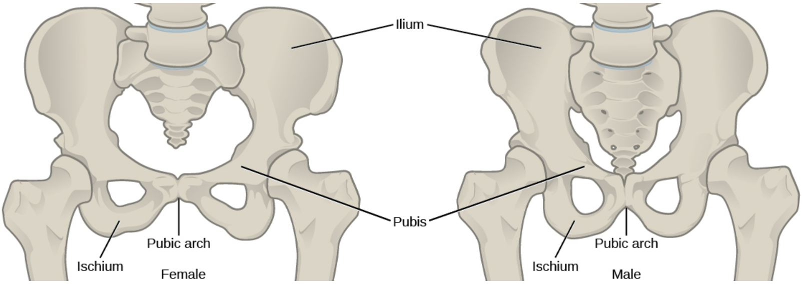
with the humerus at the elbow. The **ulna** is located on the medial aspect (pinky-finger side) of the forearm. It is longer than the radius. The ulna articulates with the humerus at the elbow. The radius and ulna also articulate with the carpal bones and with each other, which in vertebrates enables a variable degree of rotation of the carpus with respect to the long axis of the limb. The hand includes the eight bones of the **carpus** (wrist), the five bones of the **metacarpus** (palm), and the 14 bones of the **phalanges** (digits). Each digit consists of three phalanges, except for the thumb, when present, which has only two.

The Pelvic Girdle

The **pelvic girdle** attaches to the lower limbs of the axial skeleton. Because it is responsible for bearing the weight of the body and for locomotion, the pelvic girdle is securely attached to the axial skeleton by strong ligaments. It also has deep sockets with robust ligaments to securely attach the femur to the body. The pelvic girdle is further strengthened by two large hip bones. In adults, the hip bones, or **coxal bones**, are formed by the fusion of three pairs of bones: the ilium, ischium, and pubis. The pelvis joins together in the anterior of the body at a joint called the pubic symphysis and with the bones of the sacrum at the posterior of the body.

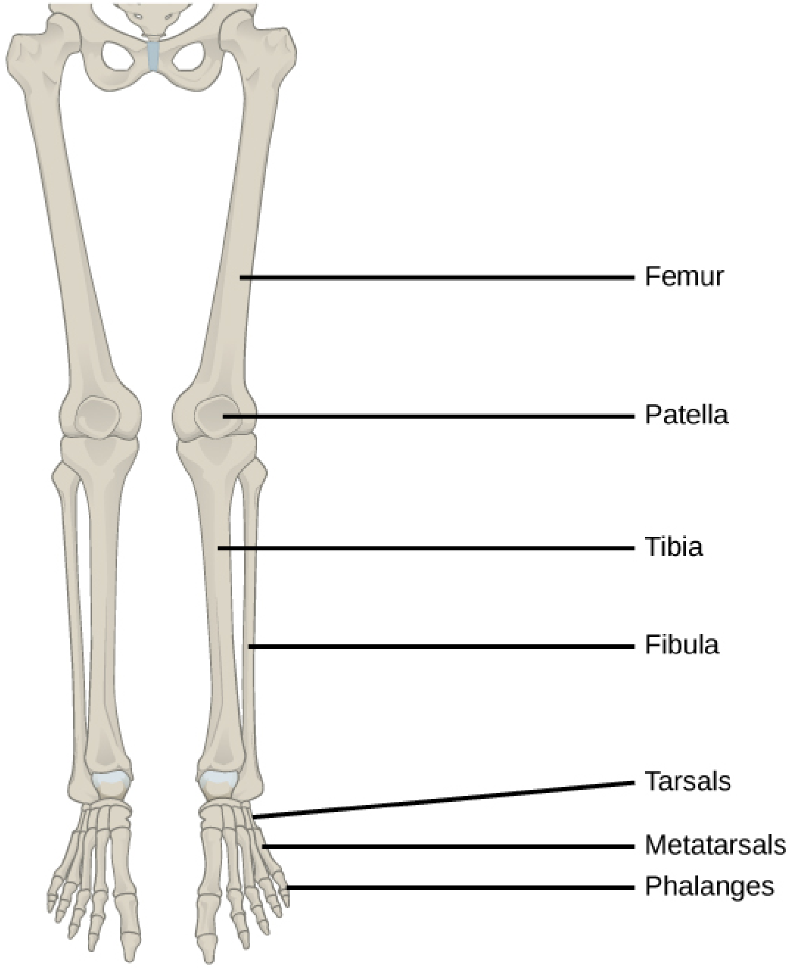
The female pelvis is slightly different from the male

pelvis. Over generations of evolution, females with a wider pubic angle and larger diameter pelvic canal reproduced more successfully. Therefore, their offspring also had pelvic anatomy that enabled successful childbirth ([link]).



The Lower Limb

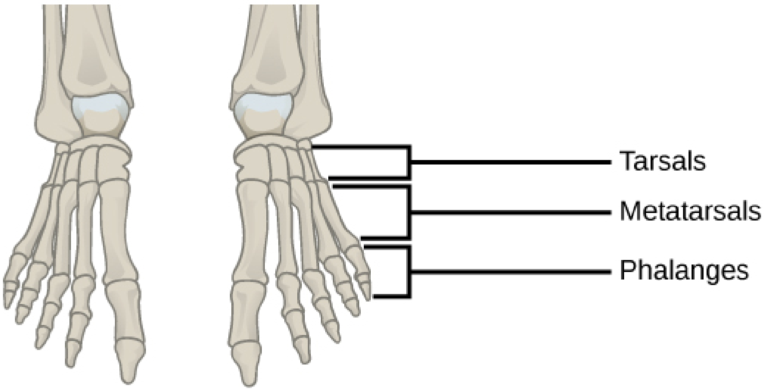
The **lower limb** consists of the thigh, the leg, and the foot. The bones of the lower limb are the femur (thigh bone), patella (kneecap), tibia and fibula (bones of the leg), tarsals (bones of the ankle), and metatarsals and phalanges (bones of the foot) ([link]). The bones of the lower limbs are thicker and stronger than the bones of the upper limbs because of the need to support the entire weight of the body and the resulting forces from locomotion. In addition to evolutionary fitness, the bones of an individual will respond to forces exerted upon them.



The **femur**, or thighbone, is the longest, heaviest, and strongest bone in the body. The femur and pelvis form the hip joint at the proximal end. At the distal end, the femur, tibia, and patella form the knee joint. The **patella**, or kneecap, is a triangular bone that lies anterior to the knee joint. The patella is embedded in the tendon of the femoral extensors (quadriceps). It improves knee extension by reducing friction. The **tibia**, or shinbone, is a large

bone of the leg that is located directly below the knee. The tibia articulates with the femur at its proximal end, with the fibula and the tarsal bones at its distal end. It is the second largest bone in the human body and is responsible for transmitting the weight of the body from the femur to the foot. The **fibula**, or calf bone, parallels and articulates with the tibia. It does not articulate with the femur and does not bear weight. The fibula acts as a site for muscle attachment and forms the lateral part of the ankle joint.

The **tarsals** are the seven bones of the ankle. The ankle transmits the weight of the body from the tibia and the fibula to the foot. The **metatarsals** are the five bones of the foot. The phalanges are the 14 bones of the toes. Each toe consists of three phalanges, except for the big toe that has only two ([link]). Variations exist in other species; for example, the horse's metacarpals and metatarsals are oriented vertically and do not make contact with the substrate.



Evolution Connection

Evolution of Body Design for Locomotion on Land

The transition of vertebrates onto land required a number of changes in body design, as movement on land presents a number of challenges for animals that are adapted to movement in water. The buoyancy of water provides a certain amount of lift, and a common form of movement by fish is lateral undulations of the entire body. This back and forth movement pushes the body against the water, creating forward movement. In most fish, the muscles of paired fins attach to girdles within the body, allowing for some control of locomotion. As certain fish began moving onto land, they retained their lateral undulation form of locomotion (anguilliform). However, instead of pushing against water, their fins or flippers became points of contact with the ground, around which they rotated their bodies.

The effect of gravity and the lack of buoyancy on land meant that body weight was suspended on the limbs, leading to increased strengthening and ossification of the limbs. The effect of gravity also required changes to the axial skeleton. Lateral undulations of land animal vertebral columns cause torsional strain. A firmer, more ossified vertebral column became common in terrestrial tetrapods because it reduces strain while providing the strength needed to support the body’s weight. In later tetrapods, the vertebrae began allowing for

vertical motion rather than lateral flexion. Another change in the axial skeleton was the loss of a direct attachment between the pectoral girdle and the head. This reduced the jarring to the head caused by the impact of the limbs on the ground. The vertebrae of the neck also evolved to allow movement of the head independently of the body. The appendicular skeleton of land animals is also different from aquatic animals. The shoulders attach to the pectoral girdle through muscles and connective tissue, thus reducing the jarring of the skull. Because of a lateral undulating vertebral column, in early tetrapods, the limbs were splayed out to the side and movement occurred by performing "push-ups." The vertebrae of these animals had to move side-to-side in a similar manner to fish and reptiles. This type of motion requires large muscles to move the limbs toward the midline; it was almost like walking while doing push-ups, and it is not an efficient use of energy. Later tetrapods have their limbs placed under their bodies, so that each stride requires less force to move forward. This resulted in decreased adductor muscle size and an increased range of motion of the scapulae. This also restricts movement primarily to one plane, creating forward motion rather than moving the limbs upward as well as forward. The femur and humerus were also rotated, so that the ends of the limbs and digits were pointed forward, in the direction of motion, rather than out to the side. By placement underneath the

body, limbs can swing forward like a pendulum to produce a stride that is more efficient for moving over land.

Section Summary

The three types of skeleton designs are hydrostatic skeletons, exoskeletons, and endoskeletons. A hydrostatic skeleton is formed by a fluid-filled compartment held under hydrostatic pressure; movement is created by the muscles producing pressure on the fluid. An exoskeleton is a hard external skeleton that protects the outer surface of an organism and enables movement through muscles attached on the inside. An endoskeleton is an internal skeleton composed of hard, mineralized tissue that also enables movement by attachment to muscles. The human skeleton is an endoskeleton that is composed of the axial and appendicular skeleton. The axial skeleton is composed of the bones of the skull, ossicles of the ear, hyoid bone, vertebral column, and ribcage. The skull consists of eight cranial bones and 14 facial bones. Six bones make up the ossicles of the middle ear, while the hyoid bone is located in the neck under the mandible. The vertebral column contains 26 bones, and it surrounds and protects the spinal cord. The

thoracic cage consists of the sternum, ribs, thoracic vertebrae, and costal cartilages. The appendicular skeleton is made up of the limbs of the upper and lower limbs. The pectoral girdle is composed of the clavicles and the scapulae. The upper limb contains 30 bones in the arm, the forearm, and the hand. The pelvic girdle attaches the lower limbs to the axial skeleton. The lower limb includes the bones of the thigh, the leg, and the foot.

Review Questions

The forearm consists of the:

1. radius and ulna
2. radius and humerus
3. ulna and humerus
4. humerus and carpus

A

The pectoral girdle consists of the:

1. clavicle and sternum
2. sternum and scapula
3. clavicle and scapula
4. clavicle and coccyx

C

All of the following are groups of vertebrae except ______, which is a curvature.

1. thoracic
2. cervical
3. lumbar
4. pelvic

D

Which of these is a facial bone?

1. frontal
2. occipital
3. lacrimal
4. temporal

C

Which of the following is **not** a true statement comparing exoskeletons and endoskeletons?

1. Endoskeletons can support larger organisms.
2. Only endoskeletons can grow as an

organism grows.
3. Exoskeletons provide greater protection of the internal organs.
4. Exoskeletons provide less mechanical leverage.

D

Critical Thinking Questions

What are the major differences between the male pelvis and female pelvis that permit childbirth in females?

The female pelvis is tilted forward and is wider, lighter, and shallower than the male pelvis. It also has a pubic angle that is broader than the male pelvis.

What are the major differences between the pelvic girdle and the pectoral girdle that allow the pelvic girdle to bear the weight of the body?

The pelvic girdle is securely attached to the

body by strong ligaments, unlike the pectoral girdle, which is sparingly attached to the ribcage. The sockets of the pelvic girdle are deep, allowing the femur to be more stable than the pectoral girdle, which has shallow sockets for the scapula. Most tetrapods have 75 percent of their weight on the front legs because the head and neck are so heavy; the advantage of the shoulder joint is more degrees of freedom in movement.

Both hydrostatic and exoskeletons can protect internal organs from harm. Contrast the ways the skeletons perform these functions.

Hydrostatic skeletons protect internal organs from harm by cushioning them from external shock. However, these skeletons do not provide protection from external trauma. Exoskeletons are hard structures that protect the organs from damage caused by their environment. However, since they are rigid, they provide little shock absorption, so the animal will need to have other ways of cushioning its internal organs.

Scoliosis is a medical condition where the spine develops a sideways curvature. How would this change interfere with the normal function of the spine?

Normal vertebral columns are stacked in a vertical line. If the spine were to curve to the side instead this would disrupt the support and cushioning functions of the vertebrae. When the spine is out of alignment, it cannot absorb shock as well so normal activities can become painful and cause back problems later in life. The curvature also disrupts posture and structure, even disrupting lung expansion in severe cases due to changes to rib location.

Glossary

appendicular skeleton
: composed of the bones of the upper limbs, which function to grasp and manipulate objects, and the lower limbs, which permit locomotion

articulation
: any place where two bones are joined

auditory ossicle
: (also, middle ear) transduces sounds from the air into vibrations in the fluid-filled cochlea

axial skeleton
: forms the central axis of the body and includes the bones of the skull, the ossicles of the middle ear, the hyoid bone of the throat,

the vertebral column, and the thoracic cage (ribcage)

carpus
eight bones that comprise the wrist

clavicle
S-shaped bone that positions the arms laterally

coxal bone
hip bone

cranial bone
one of eight bones that form the cranial cavity that encloses the brain and serves as an attachment site for the muscles of the head and neck

endoskeleton
skeleton of living cells that produces a hard, mineralized tissue located within the soft tissue of organisms

exoskeleton
a secreted cellular product external skeleton that consists of a hard encasement on the surface of an organism

facial bone
one of the 14 bones that form the face; provides cavities for the sense organs (eyes,

mouth, and nose) and attachment points for facial muscles

femur
: (also, thighbone) longest, heaviest, and strongest bone in the body

fibula
: (also, calf bone) parallels and articulates with the tibia

forearm
: extends from the elbow to the wrist and consists of two bones: the ulna and the radius

humerus
: only bone of the arm

hydrostatic skeleton
: skeleton that consists of aqueous fluid held under pressure in a closed body compartment

hyoid bone
: lies below the mandible in the front of the neck

intervertebral disc
: composed of fibrous cartilage; lies between adjacent vertebrae from the second cervical vertebra to the sacrum

lower limb
: consists of the thigh, the leg, and the foot

metacarpus
: five bones that comprise the palm

metatarsal
: one of the five bones of the foot

patella
: (also, kneecap) triangular bone that lies anterior to the knee joint

pectoral girdle
: bones that transmit the force generated by the upper limbs to the axial skeleton

phalange
: one of the bones of the fingers or toes

pelvic girdle
: bones that transmit the force generated by the lower limbs to the axial skeleton

radius
: bone located along the lateral (thumb) side of the forearm; articulates with the humerus at the elbow

rib
: one of 12 pairs of long, curved bones that attach to the thoracic vertebrae and curve toward the front of the body to form the ribcage

scapula

flat, triangular bone located at the posterior pectoral girdle

skull
: bone that supports the structures of the face and protects the brain

sternum
: (also, breastbone) long, flat bone located at the front of the chest

tarsal
: one of the seven bones of the ankle

thoracic cage
: (also, ribcage) skeleton of the chest, which consists of the ribs, thoracic vertebrae, sternum, and costal cartilages

tibia
: (also, shinbone) large bone of the leg that is located directly below the knee

ulna
: bone located on the medial aspect (pinky-finger side) of the forearm

vertebral column
: (also, spine) surrounds and protects the spinal cord, supports the head, and acts as an attachment point for ribs and muscles of the back and neck

Bone

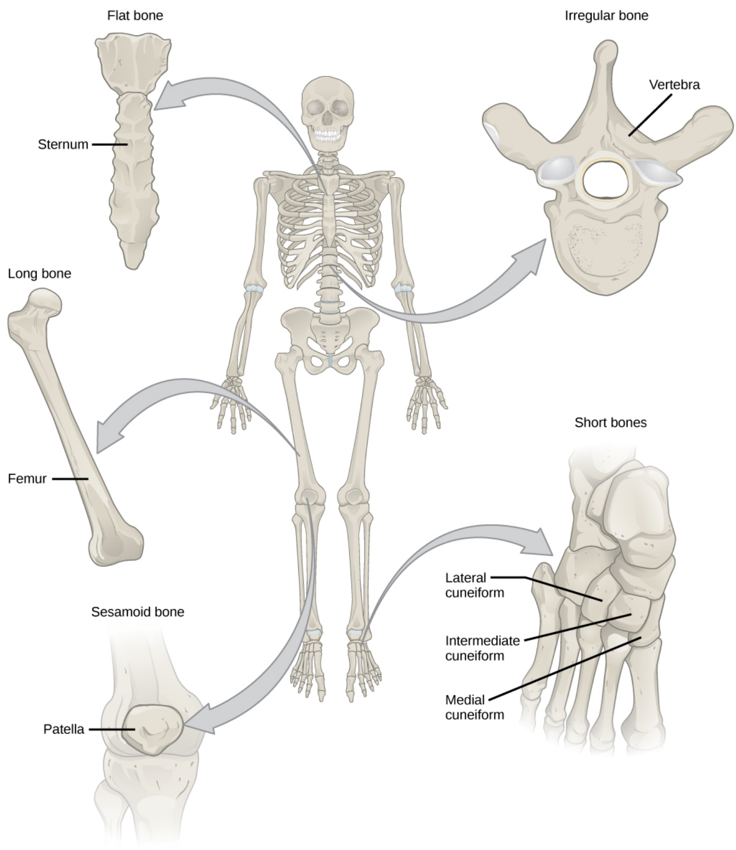
By the end of this section, you will be able to do the following:

- Classify the different types of bones in the skeleton
- Explain the role of the different cell types in bone
- Explain how bone forms during development

Bone, or **osseous tissue**, is a connective tissue that constitutes the endoskeleton. It contains specialized cells and a matrix of mineral salts and collagen fibers.

The mineral salts primarily include hydroxyapatite, a mineral formed from calcium phosphate. **Calcification** is the process of deposition of mineral salts on the collagen fiber matrix that crystallizes and hardens the tissue. The process of calcification only occurs in the presence of collagen fibers.

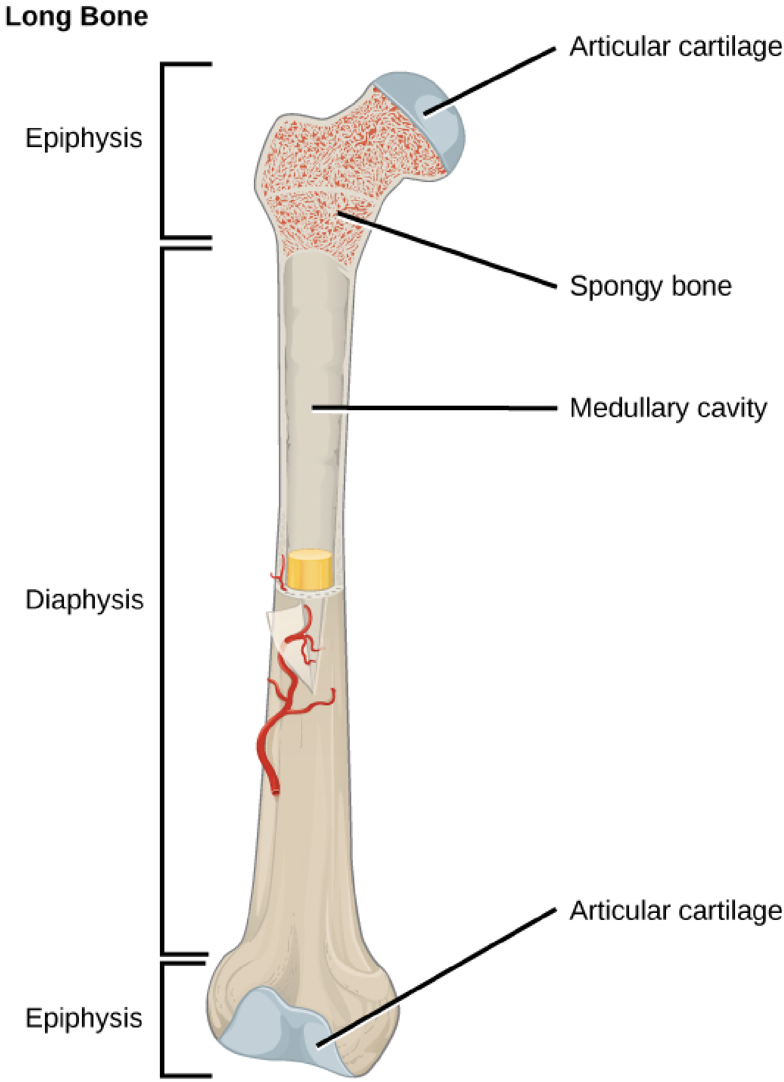
The bones of the human skeleton are classified by their shape: long bones, short bones, flat bones, sutural bones, sesamoid bones, and irregular bones ([link]).
Shown are different types of bones: flat, irregular, long, short, and sesamoid.



Long bones are longer than they are wide and have a shaft and two ends. The **diaphysis**, or central shaft, contains bone marrow in a marrow cavity. The rounded ends, the **epiphyses**, are covered with articular cartilage and are filled with red bone marrow, which produces blood cells ([link]). Most of the limb bones are long bones—for example, the femur, tibia, ulna, and radius. Exceptions to this include the patella and the bones of the wrist and ankle.

The long bone is covered by articular cartilage at

either end and contains bone marrow (shown in yellow in this illustration) in the marrow cavity.



Short bones, or cuboidal bones, are bones that are the same width and length, giving them a cube-like shape. For example, the bones of the wrist (carpals) and ankle (tarsals) are short bones ([link]).

Flat bones are thin and relatively broad bones that are found where extensive protection of organs is required or where broad surfaces of muscle attachment are required. Examples of flat bones are the sternum (breast bone), ribs, scapulae (shoulder blades), and the roof of the skull ([link]).

Irregular bones are bones with complex shapes. These bones may have short, flat, notched, or ridged surfaces. Examples of irregular bones are the vertebrae, hip bones, and several skull bones.

Sesamoid bones are small, flat bones and are shaped similarly to a sesame seed. The patellae are sesamoid bones ([link]). Sesamoid bones develop inside tendons and may be found near joints at the knees, hands, and feet.
The patella of the knee is an example of a sesamoid bone.

Sutural bones are small, flat, irregularly shaped bones. They may be found between the flat bones of the skull. They vary in number, shape, size, and position.

Trabeculae in spongy bone are arranged such that one side of the bone bears tension and the other withstands compression.

Bone Tissue

Bones are considered organs because they contain various types of tissue, such as blood, connective tissue, nerves, and bone tissue. Osteocytes, the living cells of bone tissue, form the mineral matrix of bones. There are two types of bone tissue:

compact and spongy.

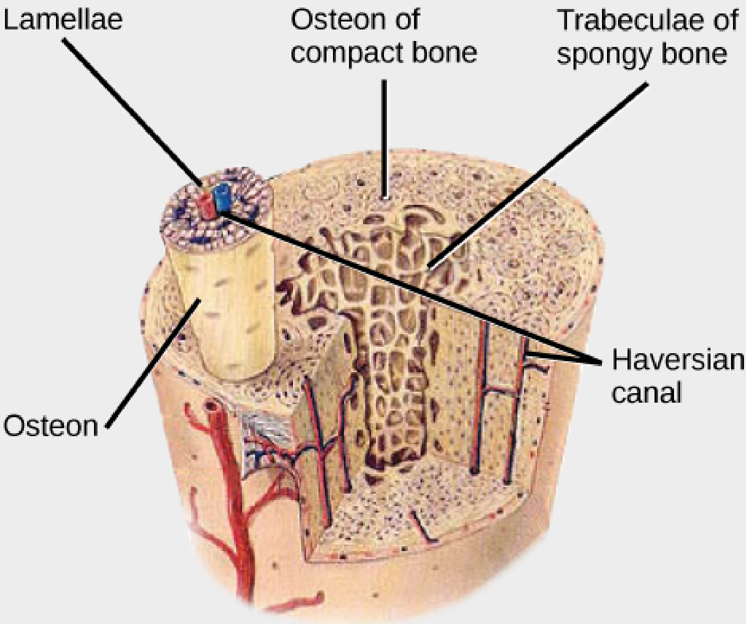
Compact Bone Tissue

Compact bone (or cortical bone) forms the hard external layer of all bones and surrounds the medullary cavity, or bone marrow. It provides protection and strength to bones. Compact bone tissue consists of units called osteons or Haversian systems. **Osteons** are cylindrical structures that contain a mineral matrix and living osteocytes connected by canaliculi, which transport blood. They are aligned parallel to the long axis of the bone. Each osteon consists of **lamellae**, which are layers of compact matrix that surround a central canal called the Haversian canal. The **Haversian canal** (osteonic canal) contains the bone's blood vessels and nerve fibers ([link]). Osteons in compact bone tissue are aligned in the same direction along lines of stress and help the bone resist bending or fracturing. Therefore, compact bone tissue is prominent in areas of bone at which stresses are applied in only a few directions.

Visual Connection

Compact bone tissue consists of osteons that are aligned parallel to the long axis of the bone, and the Haversian canal that contains the bone's blood vessels and nerve fibers. The inner layer of bones

consists of spongy bone tissue. The small dark ovals in the osteon represent the living osteocytes. (credit: modification of work by NCI, NIH)



Which of the following statements about bone tissue is false?

1. Compact bone tissue is made of cylindrical osteons that are aligned such that they travel the length of the bone.
2. Haversian canals contain blood vessels only.
3. Haversian canals contain blood vessels and nerve fibers.
4. Spongy tissue is found on the interior of the bone, and compact bone tissue is found on the exterior.

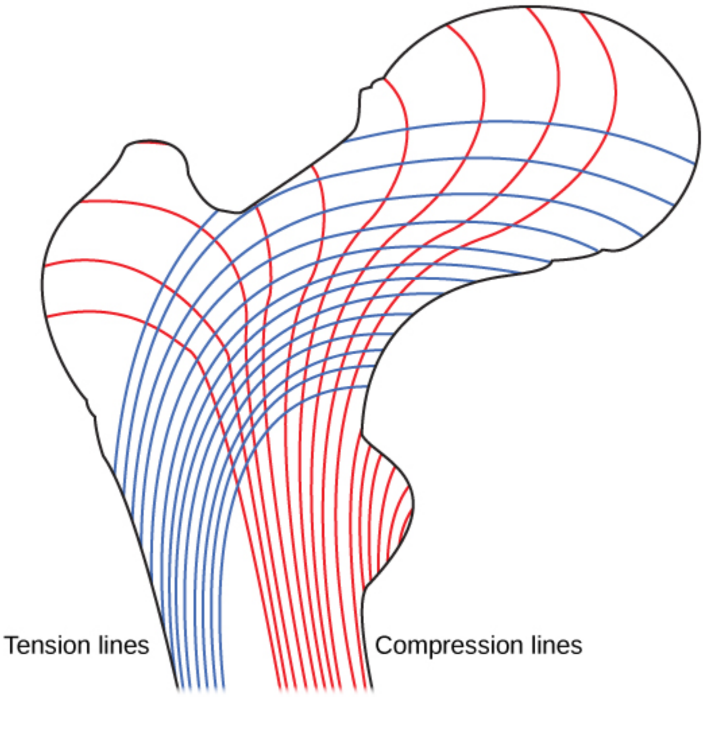
Spongy Bone Tissue

Whereas compact bone tissue forms the outer layer of all bones, **spongy bone** or cancellous bone forms the inner layer of all bones. Spongy bone tissue does not contain osteons that constitute compact bone tissue. Instead, it consists of **trabeculae**, which are lamellae that are arranged as rods or plates. Red bone marrow is found between the trabuculae. Blood vessels within this tissue deliver nutrients to osteocytes and remove waste. The red bone marrow of the femur and the interior of other large bones, such as the ileum, forms blood cells.

Spongy bone reduces the density of bone and allows the ends of long bones to compress as the result of stresses applied to the bone. Spongy bone is prominent in areas of bones that are not heavily stressed or where stresses arrive from many directions. The epiphyses of bones, such as the neck of the femur, are subject to stress from many directions. Imagine laying a heavy framed picture flat on the floor. You could hold up one side of the picture with a toothpick if the toothpick was perpendicular to the floor and the picture. Now drill a hole and stick the toothpick into the wall to hang up the picture. In this case, the function of the toothpick is to transmit the downward pressure of the picture to the wall. The force on the picture is straight down to the floor, but the force on the toothpick is both the picture wire pulling down and

the bottom of the hole in the wall pushing up. The toothpick will break off right at the wall.

The neck of the femur is horizontal like the toothpick in the wall. The weight of the body pushes it down near the joint, but the vertical diaphysis of the femur pushes it up at the other end. The neck of the femur must be strong enough to transfer the downward force of the body weight horizontally to the vertical shaft of the femur ([link]).



Link to Learning

View micrographs of musculoskeletal tissues as you review the anatomy.

Cell Types in Bones

Bone consists of four types of cells: osteoblasts, osteoclasts, osteocytes, and osteoprogenitor cells. **Osteoblasts** are bone cells that are responsible for bone formation. Osteoblasts synthesize and secrete the organic part and inorganic part of the extracellular matrix of bone tissue, and collagen fibers. Osteoblasts become trapped in these secretions and differentiate into less active osteocytes. **Osteoclasts** are large bone cells with up to 50 nuclei. They remove bone structure by releasing lysosomal enzymes and acids that dissolve the bony matrix. These minerals, released from bones into the blood, help regulate calcium concentrations in body fluids. Bone may also be resorbed for remodeling, if the applied stresses have changed. **Osteocytes** are mature bone cells and are the main cells in bony connective tissue; these cells cannot divide. Osteocytes maintain normal bone structure by recycling the mineral salts in the bony matrix. **Osteoprogenitor cells** are squamous stem cells that divide to produce daughter cells that differentiate into osteoblasts. Osteoprogenitor cells are important in the repair of fractures.

Development of Bone

Ossification, or osteogenesis, is the process of bone formation by osteoblasts. Ossification is distinct from the process of calcification; whereas calcification takes place during the ossification of bones, it can also occur in other tissues. Ossification begins approximately six weeks after fertilization in an embryo. Before this time, the embryonic skeleton consists entirely of fibrous membranes and hyaline cartilage. The development of bone from fibrous membranes is called intramembranous ossification; development from hyaline cartilage is called endochondral ossification. Bone growth continues until approximately age 25. Bones can grow in thickness throughout life, but after age 25, ossification functions primarily in bone remodeling and repair.

Endochondral ossification is the process of bone development from hyaline cartilage. The periosteum is the connective tissue on the outside of bone that acts as the interface between bone, blood vessels, tendons, and ligaments.

Intramembranous Ossification

Intramembranous ossification is the process of bone development from fibrous membranes. It is involved in the formation of the flat bones of the skull, the mandible, and the clavicles. Ossification begins as mesenchymal cells form a template of the future bone. They then differentiate into osteoblasts at the ossification center. Osteoblasts secrete the

extracellular matrix and deposit calcium, which hardens the matrix. The non-mineralized portion of the bone or osteoid continues to form around blood vessels, forming spongy bone. Connective tissue in the matrix differentiates into red bone marrow in the fetus. The spongy bone is remodeled into a thin layer of compact bone on the surface of the spongy bone.

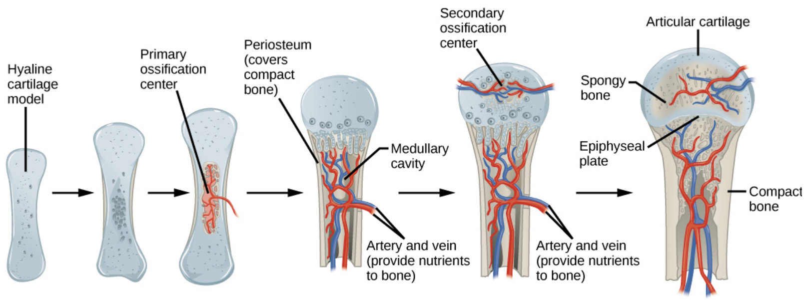
Endochondral Ossification

Endochondral ossification is the process of bone development from hyaline cartilage. All of the bones of the body, except for the flat bones of the skull, mandible, and clavicles, are formed through endochondral ossification.

In long bones, chondrocytes form a template of the hyaline cartilage diaphysis. Responding to complex developmental signals, the matrix begins to calcify. This calcification prevents diffusion of nutrients into the matrix, resulting in chondrocytes dying and the opening up of cavities in the diaphysis cartilage. Blood vessels invade the cavities, and osteoblasts and osteoclasts modify the calcified cartilage matrix into spongy bone. Osteoclasts then break down some of the spongy bone to create a marrow, or medullary, cavity in the center of the diaphysis. Dense, irregular connective tissue forms a sheath (periosteum) around the bones. The periosteum assists in attaching the bone to surrounding tissues,

tendons, and ligaments. The bone continues to grow and elongate as the cartilage cells at the epiphyses divide.

In the last stage of prenatal bone development, the centers of the epiphyses begin to calcify. Secondary ossification centers form in the epiphyses as blood vessels and osteoblasts enter these areas and convert hyaline cartilage into spongy bone. Until adolescence, hyaline cartilage persists at the **epiphyseal plate** (growth plate), which is the region between the diaphysis and epiphysis that is responsible for the lengthwise growth of long bones ([link]).



Growth of Bone

Long bones continue to lengthen, potentially until adolescence, through the addition of bone tissue at the epiphyseal plate. They also increase in width through appositional growth.

Lengthening of Long Bones

Chondrocytes on the epiphyseal side of the epiphyseal plate divide; one cell remains undifferentiated near the epiphysis, and one cell moves toward the diaphysis. The cells, which are pushed from the epiphysis, mature and are destroyed by calcification. This process replaces cartilage with bone on the diaphyseal side of the plate, resulting in a lengthening of the bone.

Long bones stop growing at around the age of 18 in females and the age of 21 in males in a process called epiphyseal plate closure. During this process, cartilage cells stop dividing and all of the cartilage is replaced by bone. The epiphyseal plate fades, leaving a structure called the epiphyseal line or epiphyseal remnant, and the epiphysis and diaphysis fuse.

Thickening of Long Bones

Appositional growth is the increase in the diameter of bones by the addition of bony tissue at the surface of bones. Osteoblasts at the bone surface secrete bone matrix, and osteoclasts on the inner surface break down bone. The osteoblasts differentiate into osteocytes. A balance between these two processes allows the bone to thicken without becoming too heavy.

After this bone is set, a callus will knit the two ends together. (credit: Bill Rhodes)

Bone Remodeling and Repair

Bone renewal continues after birth into adulthood. **Bone remodeling** is the replacement of old bone tissue by new bone tissue. It involves the processes of bone deposition by osteoblasts and bone resorption by osteoclasts. Normal bone growth requires vitamins D, C, and A, plus minerals such as calcium, phosphorous, and magnesium. Hormones such as parathyroid hormone, growth hormone, and calcitonin are also required for proper bone growth and maintenance.

Bone turnover rates are quite high, with five to seven percent of bone mass being recycled every week. Differences in turnover rate exist in different areas of the skeleton and in different areas of a bone. For example, the bone in the head of the femur may be fully replaced every six months, whereas the bone along the shaft is altered much more slowly.

Bone remodeling allows bones to adapt to stresses by becoming thicker and stronger when subjected to stress. Bones that are not subject to normal stress, for example when a limb is in a cast, will begin to lose mass. A fractured or broken bone undergoes repair through four stages:

1. Blood vessels in the broken bone tear and hemorrhage, resulting in the formation of

clotted blood, or a hematoma, at the site of the break. The severed blood vessels at the broken ends of the bone are sealed by the clotting process, and bone cells that are deprived of nutrients begin to die.
2. Within days of the fracture, capillaries grow into the hematoma, and phagocytic cells begin to clear away the dead cells. Though fragments of the blood clot may remain, fibroblasts and osteoblasts enter the area and begin to reform bone. Fibroblasts produce collagen fibers that connect the broken bone ends, and osteoblasts start to form spongy bone. The repair tissue between the broken bone ends is called the fibrocartilaginous callus, as it is composed of both hyaline and fibrocartilage ([link]). Some bone spicules may also appear at this point.
3. The fibrocartilaginous callus is converted into a bony callus of spongy bone. It takes about two months for the broken bone ends to be firmly joined together after the fracture. This is similar to the endochondral formation of bone, as cartilage becomes ossified; osteoblasts, osteoclasts, and bone matrix are present.
4. The bony callus is then remodelled by osteoclasts and osteoblasts, with excess material on the exterior of the bone and within the medullary cavity being removed. Compact bone is added to create bone tissue that is similar to the original, unbroken bone. This remodeling can take many months, and the

bone may remain uneven for years.

Scientific Method Connection

Decalcification of Bones

Question: What effect does the removal of calcium and collagen have on bone structure?

Background: Conduct a literature search on the role of calcium and collagen in maintaining bone structure. Conduct a literature search on diseases in which bone structure is compromised.

Hypothesis: Develop a hypothesis that states predictions of the flexibility, strength, and mass of bones that have had the calcium and collagen components removed. Develop a hypothesis

regarding the attempt to add calcium back to decalcified bones.

Test the hypothesis: Test the prediction by removing calcium from chicken bones by placing them in a jar of vinegar for seven days. Test the hypothesis regarding adding calcium back to decalcified bone by placing the decalcified chicken bones into a jar of water with calcium supplements added. Test the prediction by denaturing the collagen from the bones by baking them at 250°C for three hours.

Analyze the data: Create a table showing the changes in bone flexibility, strength, and mass in the three different environments.

Report the results: Under which conditions was the bone most flexible? Under which conditions was the bone the strongest?

Draw a conclusion: Did the results support or refute the hypothesis? How do the results observed in this experiment correspond to diseases that destroy bone tissue?

Section Summary

Bone, or osseous tissue, is connective tissue that includes specialized cells, mineral salts, and collagen fibers. The human skeleton can be divided

into long bones, short bones, flat bones, and irregular bones. Compact bone tissue is composed of osteons and forms the external layer of all bones. Spongy bone tissue is composed of trabeculae and forms the inner part of all bones. Four types of cells compose bony tissue: osteocytes, osteoclasts, osteoprogenitor cells, and osteoblasts. Ossification is the process of bone formation by osteoblasts. Intramembranous ossification is the process of bone development from fibrous membranes. Endochondral ossification is the process of bone development from hyaline cartilage. Long bones lengthen as chondrocytes divide and secrete hyaline cartilage. Osteoblasts replace cartilage with bone. Appositional growth is the increase in the diameter of bones by the addition of bone tissue at the surface of bones. Bone remodeling involves the processes of bone deposition by osteoblasts and bone resorption by osteoclasts. Bone repair occurs in four stages and can take several months.

Art Exercise

[link] Which of the following statements about bone tissue is false?

1. Compact bone tissue is made of cylindrical osteons that are aligned such that they

travel the length of the bone.
2. Haversian canals contain blood vessels only.
3. Haversian canals contain blood vessels and nerve fibers.
4. Spongy tissue is found on the interior of the bone, and compact bone tissue is found on the exterior.

[link]B

Review Questions

The Haversian canal:

1. is arranged as rods or plates
2. contains the bone’s blood vessels and nerve fibers
3. is responsible for the lengthwise growth of long bones
4. synthesizes and secretes matrix

B

The epiphyseal plate:

1. is arranged as rods or plates
2. contains the bone’s blood vessels and nerve fibers
3. is responsible for the lengthwise growth of long bones
4. synthesizes and secretes bone matrix

C

The cells responsible for bone resorption are _____.

1. osteoclasts
2. osteoblasts
3. fibroblasts
4. osteocytes

A

Compact bone is composed of _____.

1. trabeculae
2. compacted collagen
3. osteons
4. calcium phosphate only

C

Osteoporosis is a condition where bones become weak and brittle. It is caused by an imbalance in the activity of which cells?

1. osteoclasts and osteoblasts
2. osteoclasts and osteocytes
3. osteoblasts and chondrocytes
4. osteocytes and chondrocytes

A

While assembling a skeleton of a new species, a scientist points to one of the bones and observes that it looks like the most likely site of leg muscle attachment. What kind of bone did she indicate?

1. sesamoid bone
2. long bone
3. trabecular bone
4. flat bone

D

Critical Thinking Questions

What are the major differences between spongy bone and compact bone?

Compact bone tissue forms the hard external layer of all bones and consists of osteons. Compact bone tissue is prominent in areas of bone at which stresses are applied in only a few directions. Spongy bone tissue forms the inner layer of all bones and consists of trabeculae. Spongy bone is prominent in areas of bones that are not heavily stressed or at which stresses arrive from many directions.

What are the roles of osteoblasts, osteocytes, and osteoclasts?

Osteocytes function in the exchange of nutrients and wastes with the blood. They also maintain normal bone structure by recycling the mineral salts in the bony matrix. Osteoclasts remove bone tissue by releasing lysosomal enzymes and acids that dissolve the bony matrix. Osteoblasts are bone cells that are responsible for bone formation.

Thalidomide was a morning sickness drug given to women that caused babies to be born without arm bones. If recent studies have

shown that thalidomide prevents the formation of new blood vessels, describe the type of bone development inhibited by the drug and what stage of ossification was affected.

Thalidomide effected the development of the long bones of the arms, disrupting endochondral ossification. The bones would have been able to develop into a template made of the calcified cartilage matrix, but new blood vessels could not be created. Since no vessels invade the template, the structure is not converted into trabecular bone.

Glossary

appositional growth
: increase in the diameter of bones by the addition of bone tissue at the surface of bones

bone
: (also, osseous tissue) connective tissue that constitutes the endoskeleton

bone remodeling
: replacement of old bone tissue by new bone tissue

calcification
: process of deposition of mineral salts in the

collagen fiber matrix that crystallizes and hardens the tissue

compact bone
: forms the hard external layer of all bones

diaphysis
: central shaft of bone, contains bone marrow in a marrow cavity

endochondral ossification
: process of bone development from hyaline cartilage

epiphyseal plate
: region between the diaphysis and epiphysis that is responsible for the lengthwise growth of long bones

epiphysis
: rounded end of bone, covered with articular cartilage and filled with red bone marrow, which produces blood cells

flat bone
: thin and relatively broad bone found where extensive protection of organs is required or where broad surfaces of muscle attachment are required

Haversian canal
: contains the bone's blood vessels and nerve

fibers

intramembranous ossification
: process of bone development from fibrous membranes

irregular bone
: bone with complex shapes; examples include vertebrae and hip bones

lamella
: layer of compact tissue that surrounds a central canal called the Haversian canal

long bone
: bone that is longer than wide, and has a shaft and two ends

osteoblast
: bone cell responsible for bone formation

osteoclast
: large bone cells with up to 50 nuclei, responsible for bone remodeling

osteocyte
: mature bone cells and the main cell in bone tissue

osseous tissue
: connective tissue that constitutes the endoskeleton

ossification
: (also, osteogenesis) process of bone formation by osteoblasts

osteon
: cylindrical structure aligned parallel to the long axis of the bone

resorption
: process by which osteoclasts release minerals stored in bones

sesamoid bone
: small, flat bone shaped like a sesame seed; develops inside tendons

short bone
: bone that has the same width and length, giving it a cube-like shape

spongy bone tissue
: forms the inner layer of all bones

sutural bone
: small, flat, irregularly shaped bone that forms between the flat bones of the cranium

trabeculae
: lamellae that are arranged as rods or plates

Joints and Skeletal Movement
By the end of this section, you will be able to do the following:

- Classify the different types of joints on the basis of structure
- Explain the role of joints in skeletal movement

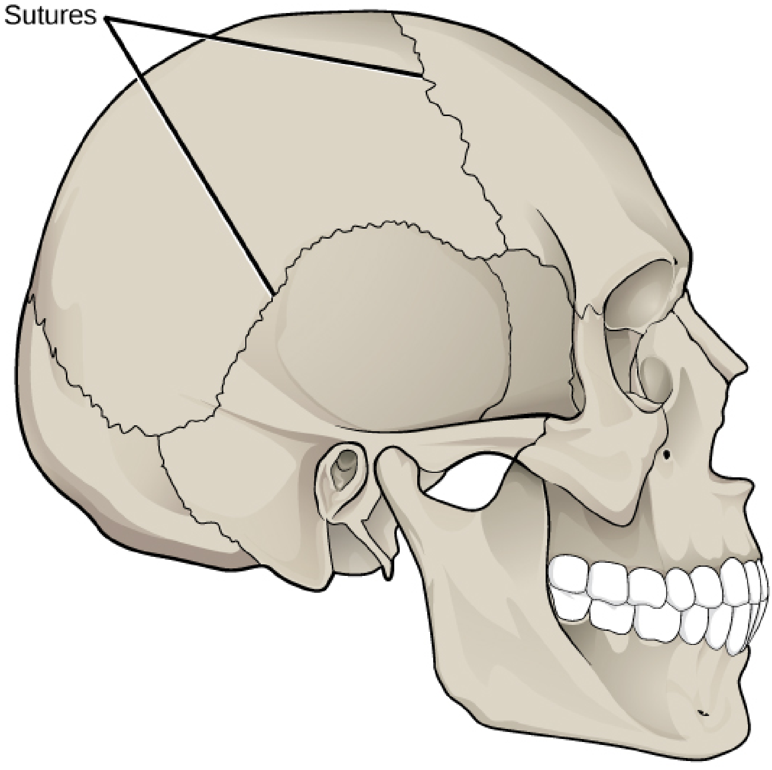
The point at which two or more bones meet is called a **joint**, or **articulation**. Joints are responsible for movement, such as the movement of limbs, and stability, such as the stability found in the bones of the skull.
Sutures are fibrous joints found only in the skull. Gomphoses are fibrous joints between the teeth and their sockets. (credit: modification of work by Gray's Anatomy) Synovial joints are the only joints that have a space or "synovial cavity" in the joint.

Classification of Joints on the Basis of Structure

There are two ways to classify joints: on the basis of their structure or on the basis of their function. The structural classification divides joints into bony, fibrous, cartilaginous, and synovial joints depending on the material composing the joint and the presence or absence of a cavity in the joint.

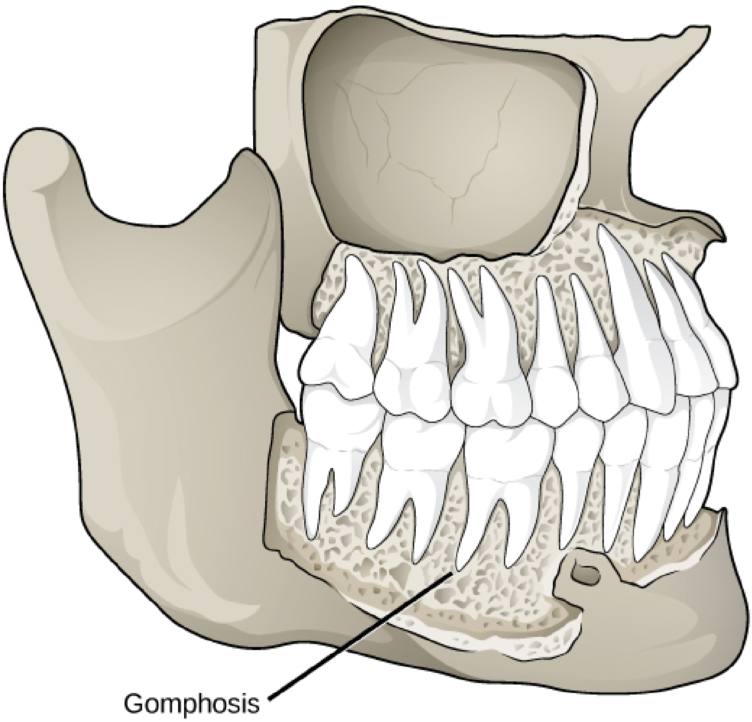
Fibrous Joints

The bones of **fibrous joints** are held together by fibrous connective tissue. There is no cavity, or space, present between the bones and so most fibrous joints do not move at all, or are only capable of minor movements. There are three types of fibrous joints: sutures, syndesmoses, and gomphoses. **Sutures** are found only in the skull and possess short fibers of connective tissue that hold the skull bones tightly in place ([link]).



Syndesmoses are joints in which the bones are connected by a band of connective tissue, allowing

for more movement than in a suture. An example of a syndesmosis is the joint of the tibia and fibula in the ankle. The amount of movement in these types of joints is determined by the length of the connective tissue fibers. **Gomphoses** occur between teeth and their sockets; the term refers to the way the tooth fits into the socket like a peg ([link]). The tooth is connected to the socket by a connective tissue referred to as the periodontal ligament.



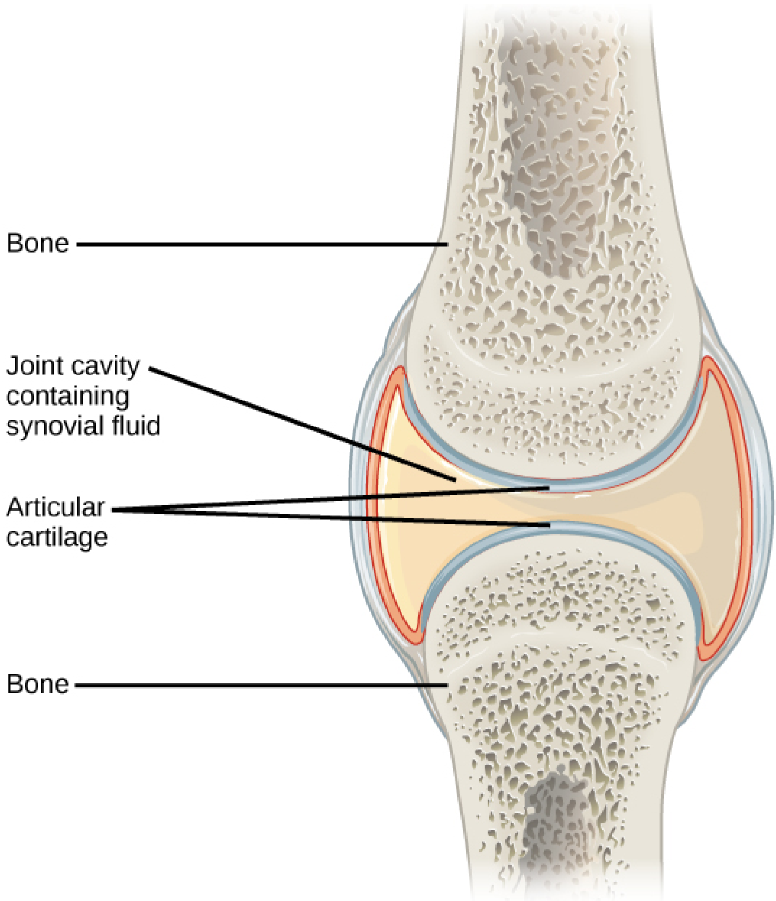
Cartilaginous Joints

Cartilaginous joints are joints in which the bones are connected by cartilage. There are two types of cartilaginous joints: synchondroses and symphyses.

In a **synchondrosis**, the bones are joined by hyaline cartilage. Synchondroses are found in the epiphyseal plates of growing bones in children. In **symphyses**, hyaline cartilage covers the end of the bone but the connection between bones occurs through fibrocartilage. Symphyses are found at the joints between vertebrae. Either type of cartilaginous joint allows for very little movement.

Synovial Joints

Synovial joints are the only joints that have a space between the adjoining bones ([link]). This space is referred to as the synovial (or joint) cavity and is filled with synovial fluid. Synovial fluid lubricates the joint, reducing friction between the bones and allowing for greater movement. The ends of the bones are covered with articular cartilage, a hyaline cartilage, and the entire joint is surrounded by an articular capsule composed of connective tissue that allows movement of the joint while resisting dislocation. Articular capsules may also possess ligaments that hold the bones together. Synovial joints are capable of the greatest movement of the three structural joint types; however, the more mobile a joint, the weaker the joint. Knees, elbows, and shoulders are examples of synovial joints.



Classification of Joints on the Basis of Function

The functional classification divides joints into three categories: synarthroses, amphiarthroses, and diarthroses. A **synarthrosis** is a joint that is immovable. This includes sutures, gomphoses, and synchondroses. **Amphiarthroses** are joints that

allow slight movement, including syndesmoses and symphyses. **Diarthroses** are joints that allow for free movement of the joint, as in synovial joints.

Movement at Synovial Joints

The wide range of movement allowed by synovial joints produces different types of movements. The movement of synovial joints can be classified as one of four different types: gliding, angular, rotational, or special movement.

Gliding Movement

Gliding movements occur as relatively flat bone surfaces move past each other. Gliding movements produce very little rotation or angular movement of the bones. The joints of the carpal and tarsal bones are examples of joints that produce gliding movements.

Angular Movement

Angular movements are produced when the angle between the bones of a joint changes. There are several different types of angular movements, including flexion, extension, hyperextension, abduction, adduction, and circumduction. **Flexion**, or bending, occurs when the angle between the

bones decreases. Moving the forearm upward at the elbow or moving the wrist to move the hand toward the forearm are examples of flexion. **Extension** is the opposite of flexion in that the angle between the bones of a joint increases. Straightening a limb after flexion is an example of extension. Extension past the regular anatomical position is referred to as **hyperextension**. This includes moving the neck back to look upward, or bending the wrist so that the hand moves away from the forearm.

Abduction occurs when a bone moves away from the midline of the body. Examples of abduction are moving the arms or legs laterally to lift them straight out to the side. **Adduction** is the movement of a bone toward the midline of the body. Movement of the limbs inward after abduction is an example of adduction. **Circumduction** is the movement of a limb in a circular motion, as in moving the arm in a circular motion.

Rotational Movement

Rotational movement is the movement of a bone as it rotates around its longitudinal axis. Rotation can be toward the midline of the body, which is referred to as **medial rotation**, or away from the midline of the body, which is referred to as **lateral rotation**. Movement of the head from side to side is an example of rotation.

Special Movements

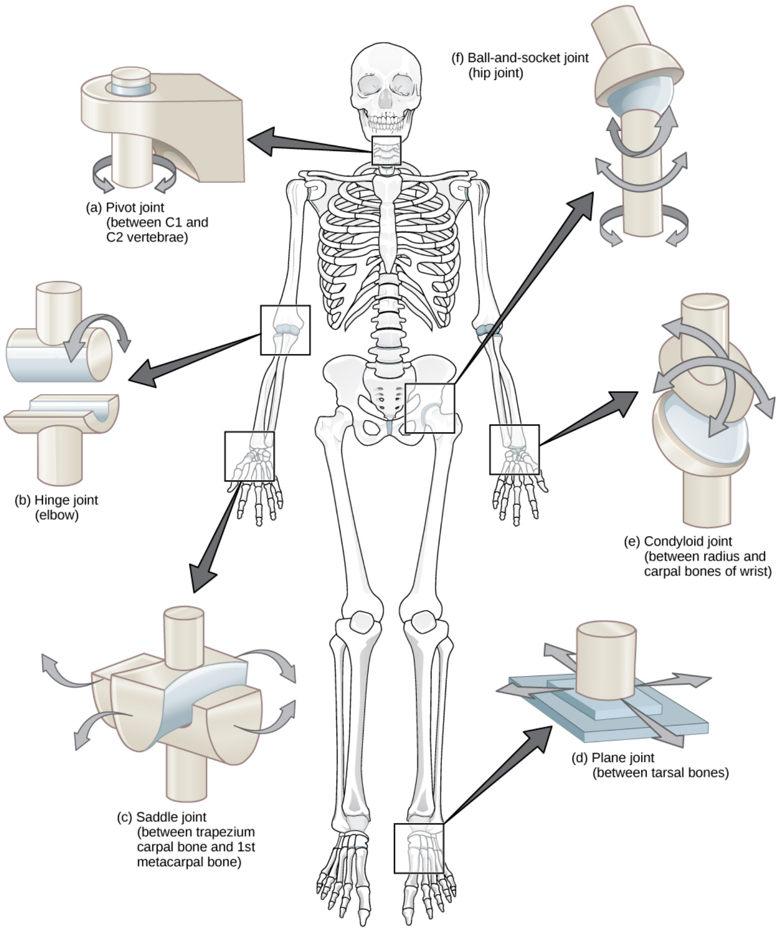
Some movements that cannot be classified as gliding, angular, or rotational are called special movements. **Inversion** involves the soles of the feet moving inward, toward the midline of the body. **Eversion** is the opposite of inversion, movement of the sole of the foot outward, away from the midline of the body. **Protraction** is the anterior movement of a bone in the horizontal plane. **Retraction** occurs as a joint moves back into position after protraction. Protraction and retraction can be seen in the movement of the mandible as the jaw is thrust outwards and then back inwards. **Elevation** is the movement of a bone upward, such as when the shoulders are shrugged, lifting the scapulae. **Depression** is the opposite of elevation—movement downward of a bone, such as after the shoulders are shrugged and the scapulae return to their normal position from an elevated position. **Dorsiflexion** is a bending at the ankle such that the toes are lifted toward the knee. **Plantar flexion** is a bending at the ankle when the heel is lifted, such as when standing on the toes. **Supination** is the movement of the radius and ulna bones of the forearm so that the palm faces forward. **Pronation** is the opposite movement, in which the palm faces backward. **Opposition** is the movement of the thumb toward the fingers of the same hand, making it possible to grasp and hold objects.

Different types of joints allow different types of

movement. Planar, hinge, pivot, condyloid, saddle, and ball-and-socket are all types of synovial joints. The joints of the carpal bones in the wrist are examples of planar joints. (credit: modification of work by Brian C. Goss) The elbow joint, where the radius articulates with the humerus, is an example of a hinge joint. (credit: modification of work by Brian C. Goss) The joint in the neck that allows the head to move back and forth is an example of a pivot joint. The metacarpophalangeal joints in the finger are examples of condyloid joints. (credit: modification of work by Gray's Anatomy) The carpometacarpal joints in the thumb are examples of saddle joints. (credit: modification of work by Brian C. Goss) The shoulder joint is an example of a ball-and-socket joint.

Types of Synovial Joints

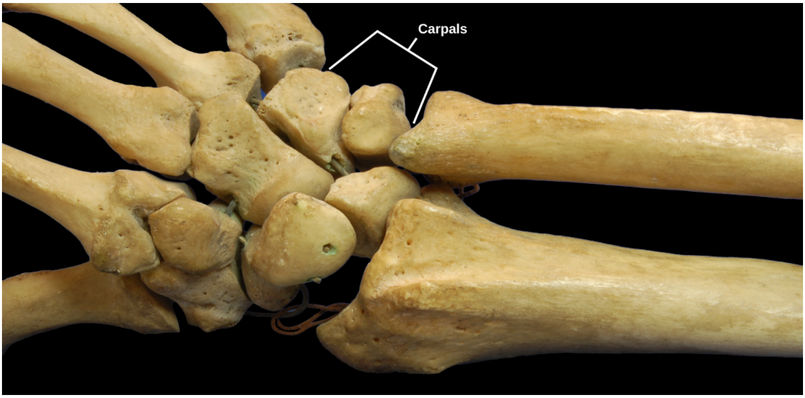
Synovial joints are further classified into six different categories on the basis of the shape and structure of the joint. The shape of the joint affects the type of movement permitted by the joint ([link]). These joints can be described as planar, hinge, pivot, condyloid, saddle, or ball-and-socket joints.



Planar Joints

Planar joints have bones with articulating surfaces that are flat or slightly curved faces. These joints allow for gliding movements, and so the joints are sometimes referred to as gliding joints. The range of motion is limited in these joints and does not involve rotation. Planar joints are found in the carpal bones in the hand and the tarsal bones of the

foot, as well as between vertebrae ([link]).



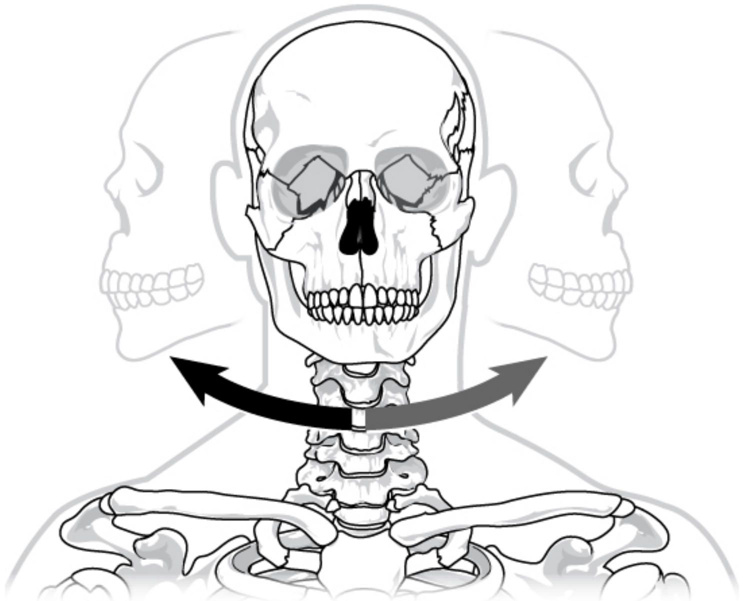
Hinge Joints

In **hinge joints**, the slightly rounded end of one bone fits into the slightly hollow end of the other bone. In this way, one bone moves while the other remains stationary, like the hinge of a door. The elbow is an example of a hinge joint. The knee is sometimes classified as a modified hinge joint ([link]).

Pivot Joints

Pivot joints consist of the rounded end of one bone fitting into a ring formed by the other bone. This

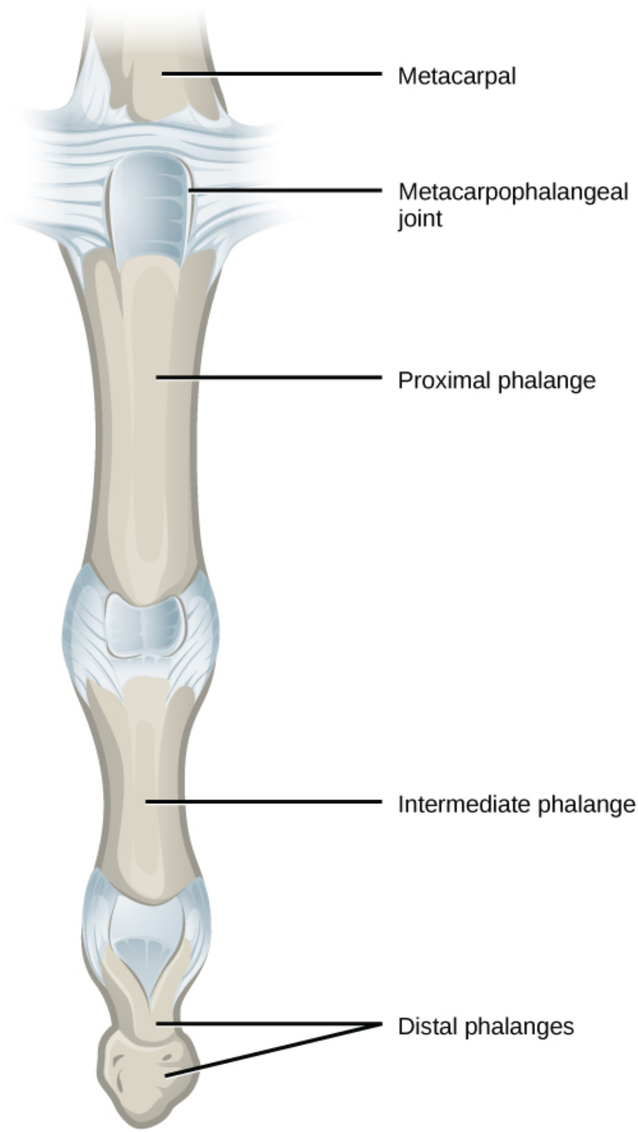
structure allows rotational movement, as the rounded bone moves around its own axis. An example of a pivot joint is the joint of the first and second vertebrae of the neck that allows the head to move back and forth ([link]). The joint of the wrist that allows the palm of the hand to be turned up and down is also a pivot joint.



Condyloid Joints

Condyloid joints consist of an oval-shaped end of one bone fitting into a similarly oval-shaped hollow of another bone ([link]). This is also sometimes called an ellipsoidal joint. This type of joint allows angular movement along two axes, as seen in the joints of the wrist and fingers, which can move both

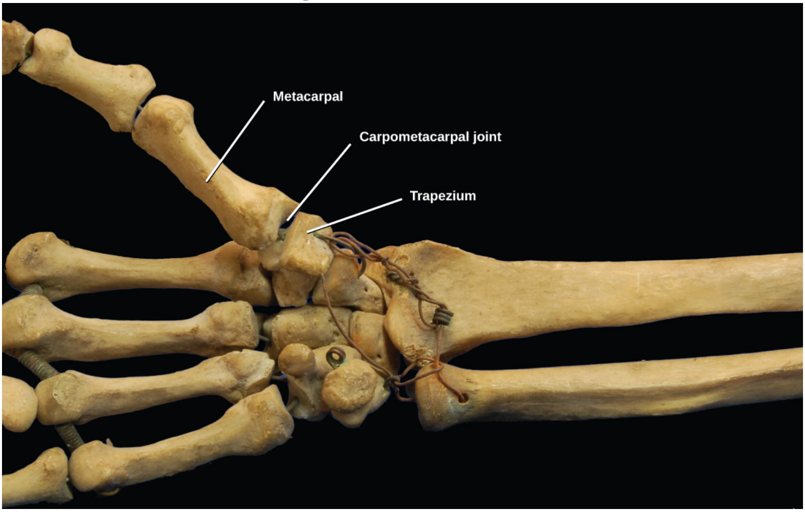
side to side and up and down.



Saddle Joints

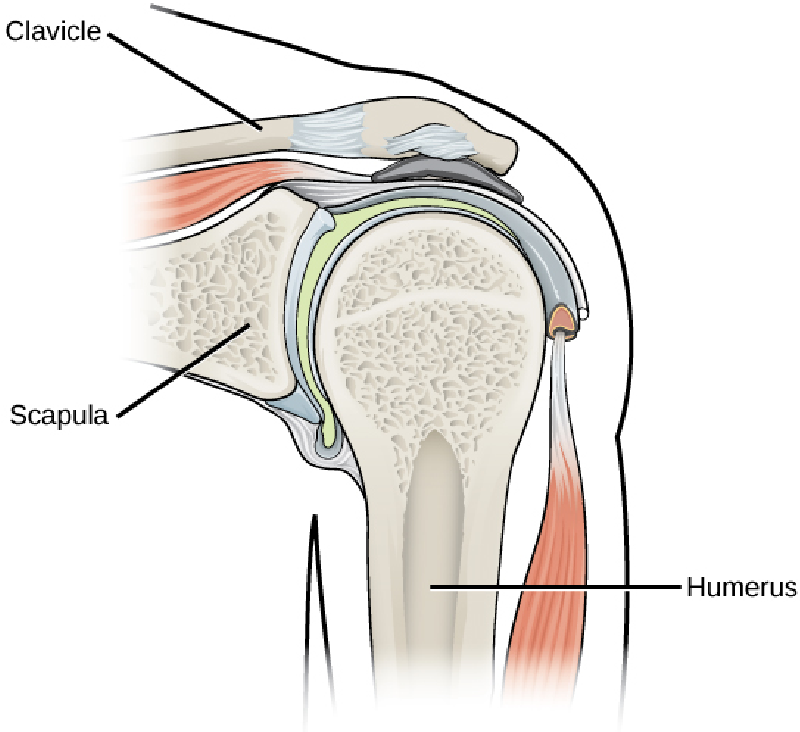
Saddle joints are so named because the ends of

each bone resemble a saddle, with concave and convex portions that fit together. Saddle joints allow angular movements similar to condyloid joints but with a greater range of motion. An example of a saddle joint is the thumb joint, which can move back and forth and up and down, but more freely than the wrist or fingers ([link]).



Ball-and-Socket Joints

Ball-and-socket joints possess a rounded, ball-like end of one bone fitting into a cuplike socket of another bone. This organization allows the greatest range of motion, as all movement types are possible in all directions. Examples of ball-and-socket joints are the shoulder and hip joints ([link]).



Link to Learning

Watch this animation showing the six types of synovial joints.
https://www.openstax.org/l/synovial_joints

Career Connection

Rheumatologist

Rheumatologists are medical doctors who specialize in the diagnosis and treatment of

disorders of the joints, muscles, and bones. They diagnose and treat diseases such as arthritis, musculoskeletal disorders, osteoporosis, and autoimmune diseases such as ankylosing spondylitis and rheumatoid arthritis.

Rheumatoid arthritis (RA) is an inflammatory disorder that primarily affects the synovial joints of the hands, feet, and cervical spine. Affected joints become swollen, stiff, and painful. Although it is known that RA is an autoimmune disease in which the body's immune system mistakenly attacks healthy tissue, the cause of RA remains unknown. Immune cells from the blood enter joints and the synovium causing cartilage breakdown, swelling, and inflammation of the joint lining. Breakdown of cartilage causes bones to rub against each other causing pain. RA is more common in women than men and the age of onset is usually 40–50 years of age.

Rheumatologists can diagnose RA on the basis of symptoms such as joint inflammation and pain, X-ray and MRI imaging, and blood tests.

Arthrography is a type of medical imaging of joints that uses a contrast agent, such as a dye, that is opaque to X-rays. This allows the soft tissue structures of joints—such as cartilage, tendons, and ligaments—to be visualized. An arthrogram differs from a regular X-ray by showing the surface of soft tissues lining the joint in addition to joint bones. An arthrogram allows early degenerative changes in joint cartilage to be detected before bones

become affected.
There is currently no cure for RA; however, rheumatologists have a number of treatment options available. Early stages can be treated with rest of the affected joints by using a cane or by using joint splints that minimize inflammation. When inflammation has decreased, exercise can be used to strengthen the muscles that surround the joint and to maintain joint flexibility. If joint damage is more extensive, medications can be used to relieve pain and decrease inflammation. Anti-inflammatory drugs such as aspirin, topical pain relievers, and corticosteroid injections may be used. Surgery may be required in cases in which joint damage is severe.

Section Summary

The structural classification of joints divides them into bony, fibrous, cartilaginous, and synovial joints. The bones of fibrous joints are held together by fibrous connective tissue; the three types of fibrous joints are sutures, syndesomes, and gomphoses. Cartilaginous joints are joints in which the bones are connected by cartilage; the two types of cartilaginous joints are synchondroses and symphyses. Synovial joints are joints that have a

space between the adjoining bones. The functional classification divides joints into three categories: synarthroses, amphiarthroses, and diarthroses. The movement of synovial joints can be classified as one of four different types: gliding, angular, rotational, or special movement. Gliding movements occur as relatively flat bone surfaces move past each other. Angular movements are produced when the angle between the bones of a joint changes. Rotational movement is the movement of a bone as it rotates around its own longitudinal axis. Special movements include inversion, eversion, protraction, retraction, elevation, depression, dorsiflexion, plantar flexion, supination, pronation, and opposition. Synovial joints are also classified into six different categories on the basis of the shape and structure of the joint: planar, hinge, pivot, condyloid, saddle, and ball-and-socket.

Review Questions

Synchondroses and symphyses are:

1. synovial joints
2. cartilaginous joints
3. fibrous joints
4. condyloid joints

B

The movement of bone away from the midline of the body is called ______.

1. circumduction
2. extension
3. adduction
4. abduction

D

Which of the following is not a characteristic of the synovial fluid?

1. lubrication
2. shock absorption
3. regulation of water balance in the joint
4. protection of articular cartilage

C

The elbow is an example of which type of joint?

1. hinge
2. pivot
3. saddle

4. gliding

A

A high ankle sprain is an injury caused by overstretching the ligaments connecting the tibia and fibula. What type of joint is involved in this sprain?

1. ball and socket
2. gomphosis
3. syndesmosis
4. symphysis

C

Critical Thinking Questions

What movements occur at the hip joint and knees as you bend down to touch your toes?

The hip joint is flexed and the knees are extended.

What movement(s) occur(s) at the scapulae when you shrug your shoulders?

Elevation is the movement of a bone upward, such as when the shoulders are shrugged, lifting the scapulae. Depression is the downward movement of a bone, such as after the shoulders are shrugged and the scapulae return to their normal position from an elevated position.

Describe the joints and motions involved in taking a step forward if a person is initially standing still. Assume the person holds his foot at the same angle throughout the motion.

Taking a step would require bending the knee (modified hinge joint) and moving the leg in the hip (ball and socket joint) since the motion of the foot is excluded. As the foot comes off the ground in the step, the hip joint is going to move the femur in a protracted motion and the knee will flex the shin toward the thigh. As the foot lands, the knee extends the leg and the hip retracts the femur.

Glossary

abduction

when a bone moves away from the midline of the body

adduction
: movement of the limbs inward after abduction

amphiarthrosis
: joint that allows slight movement; includes syndesmoses and symphyses

angular movement
: produced when the angle between the bones of a joint changes

ball-and-socket joint
: joint with a rounded, ball-like end of one bone fitting into a cuplike socket of another bone

cartilaginous joint
: joint in which the bones are connected by cartilage

circumduction
: movement of a limb in a circular motion

condyloid joint
: oval-shaped end of one bone fitting into a similarly oval-shaped hollow of another bone

depression
: movement downward of a bone, such as after

the shoulders are shrugged and the scapulae return to their normal position from an elevated position; opposite of elevation

diarthrosis
: joint that allows for free movement of the joint; found in synovial joints

dorsiflexion
: bending at the ankle such that the toes are lifted toward the knee

elevation
: movement of a bone upward, such as when the shoulders are shrugged, lifting the scapulae

eversion
: movement of the sole of the foot outward, away from the midline of the body; opposite of inversion

extension
: movement in which the angle between the bones of a joint increases; opposite of flexion

fibrous joint
: joint held together by fibrous connective tissue

flexion
: movement in which the angle between the

bones decreases; opposite of extension

gliding movement
: when relatively flat bone surfaces move past each other

gomphosis
: the joint in which the tooth fits into the socket like a peg

hinge joint
: slightly rounded end of one bone fits into the slightly hollow end of the other bone

hyperextension
: extension past the regular anatomical position

inversion
: soles of the feet moving inward, toward the midline of the body

joint
: point at which two or more bones meet

lateral rotation
: rotation away from the midline of the body

medial rotation
: rotation toward the midline of the body

opposition
: movement of the thumb toward the fingers of the same hand, making it possible to grasp

and hold objects

plantar flexion
: bending at the ankle such that the heel is lifted, such as when standing on the toes

planar joint
: joint with bones whose articulating surfaces are flat

pivot joint
: joint with the rounded end of one bone fitting into a ring formed by the other bone

pronation
: movement in which the palm faces backward

protraction
: anterior movement of a bone in the horizontal plane

retraction
: movement in which a joint moves back into position after protraction

rotational movement
: movement of a bone as it rotates around its own longitudinal axis

saddle joint
: joint with concave and convex portions that fit together; named because the ends of each bone resemble a saddle

supination
: movement of the radius and ulna bones of the forearm so that the palm faces forward

suture
: short fiber of connective tissue that holds the skull bones tightly in place; found only in the skull

synarthrosis
: joint that is immovable

symphysis
: hyaline cartilage covers the end of the bone, but the connection between bones occurs through fibrocartilage; symphyses are found at the joints between vertebrae

synchondrosis
: bones joined by hyaline cartilage; synchondroses are found in the epiphyseal plates of growing bones in children

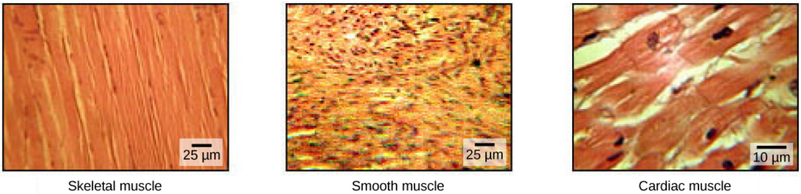
syndesmosis
: joint in which the bones are connected by a band of connective tissue, allowing for more movement than in a suture

synovial joint
: only joint that has a space between the adjoining bones

Muscle Contraction and Locomotion
By the end of this section, you will be able to do the following:

- Classify the different types of muscle tissue
- Explain the role of muscles in locomotion

Muscle cells are specialized for contraction. Muscles allow for motions such as walking, and they also facilitate bodily processes such as respiration and digestion. The body contains three types of muscle tissue: skeletal muscle, cardiac muscle, and smooth muscle ([link]).
The body contains three types of muscle tissue: skeletal muscle, smooth muscle, and cardiac muscle, visualized here using light microscopy. Smooth muscle cells are short, tapered at each end, and have only one plump nucleus in each. Cardiac muscle cells are branched and striated, but short. The cytoplasm may branch, and they have one nucleus in the center of the cell. (credit: modification of work by NCI, NIH; scale-bar data from Matt Russell)



Skeletal muscle tissue forms skeletal muscles, which attach to bones or skin and control

locomotion and any movement that can be consciously controlled. Because it can be controlled by thought, skeletal muscle is also called voluntary muscle. Skeletal muscles are long and cylindrical in appearance; when viewed under a microscope, skeletal muscle tissue has a striped or striated appearance. The striations are caused by the regular arrangement of contractile proteins (actin and myosin). **Actin** is a globular contractile protein that interacts with **myosin** for muscle contraction. Skeletal muscle also has multiple nuclei present in a single cell.

Smooth muscle tissue occurs in the walls of hollow organs such as the intestines, stomach, and urinary bladder, and around passages such as the respiratory tract and blood vessels. Smooth muscle has no striations, is not under voluntary control, has only one nucleus per cell, is tapered at both ends, and is called involuntary muscle.

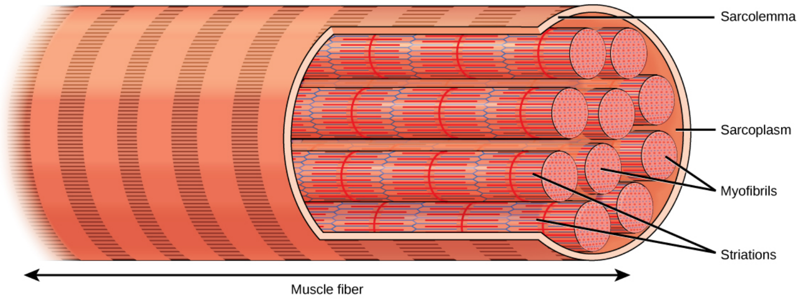
Cardiac muscle tissue is only found in the heart, and cardiac contractions pump blood throughout the body and maintain blood pressure. Like skeletal muscle, cardiac muscle is striated, but unlike skeletal muscle, cardiac muscle cannot be consciously controlled and is called involuntary muscle. It has one nucleus per cell, is branched, and is distinguished by the presence of intercalated disks.

A skeletal muscle cell is surrounded by a plasma

membrane called the sarcolemma with a cytoplasm called the sarcoplasm. A muscle fiber is composed of many fibrils, packaged into orderly units. A sarcomere is the region from one Z line to the next Z line. Many sarcomeres are present in a myofibril, resulting in the striation pattern characteristic of skeletal muscle.

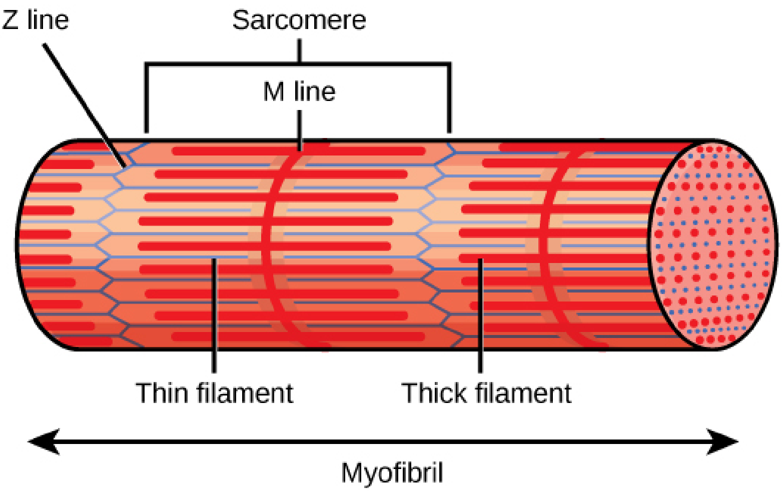
Skeletal Muscle Fiber Structure

Each skeletal muscle fiber is a skeletal muscle cell. These cells are incredibly large, with diameters of up to 100 μm and lengths of up to 30 cm. The plasma membrane of a skeletal muscle fiber is called the **sarcolemma**. The sarcolemma is the site of action potential conduction, which triggers muscle contraction. Within each muscle fiber are **myofibrils**—long cylindrical structures that lie parallel to the muscle fiber. Myofibrils run the entire length of the muscle fiber, and because they are only approximately 1.2 μm in diameter, hundreds to thousands can be found inside one muscle fiber. They attach to the sarcolemma at their ends, so that as myofibrils shorten, the entire muscle cell contracts ([link]).



The striated appearance of skeletal muscle tissue is a result of repeating bands of the proteins actin and myosin that are present along the length of myofibrils. Dark A bands and light I bands repeat along myofibrils, and the alignment of myofibrils in the cell causes the entire cell to appear striated or banded.

Each I band has a dense line running vertically through the middle called a Z disc or Z line. The Z discs mark the border of units called **sarcomeres**, which are the functional units of skeletal muscle. One sarcomere is the space between two consecutive Z discs and contains one entire A band and two halves of an I band, one on either side of the A band. A myofibril is composed of many sarcomeres running along its length, and as the sarcomeres individually contract, the myofibrils and muscle cells shorten ([link]).



Myofibrils are composed of smaller structures called **myofilaments**. There are two main types of filaments: thick filaments and thin filaments; each has different compositions and locations. **Thick filaments** occur only in the A band of a myofibril. **Thin filaments** attach to a protein in the Z disc called alpha-actinin and occur across the entire length of the I band and partway into the A band. The region at which thick and thin filaments overlap has a dense appearance, as there is little space between the filaments. Thin filaments do not extend all the way into the A bands, leaving a central region of the A band that only contains thick filaments. This central region of the A band looks slightly lighter than the rest of the A band and is called the H zone. The middle of the H zone has a vertical line called the M line, at which accessory proteins hold together thick filaments. Both the Z disc and the M line hold myofilaments in place to

maintain the structural arrangement and layering of the myofibril. Myofibrils are connected to each other by intermediate, or desmin, filaments that attach to the Z disc.

Thick and thin filaments are themselves composed of proteins. Thick filaments are composed of the protein myosin. The tail of a myosin molecule connects with other myosin molecules to form the central region of a thick filament near the M line, whereas the heads align on either side of the thick filament where the thin filaments overlap. The primary component of thin filaments is the actin protein. Two other components of the thin filament are tropomyosin and troponin. Actin has binding sites for myosin attachment. Strands of tropomyosin block the binding sites and prevent actin–myosin interactions when the muscles are at rest. Troponin consists of three globular subunits. One subunit binds to tropomyosin, one subunit binds to actin, and one subunit binds Ca^{2+} ions.

Link to Learning

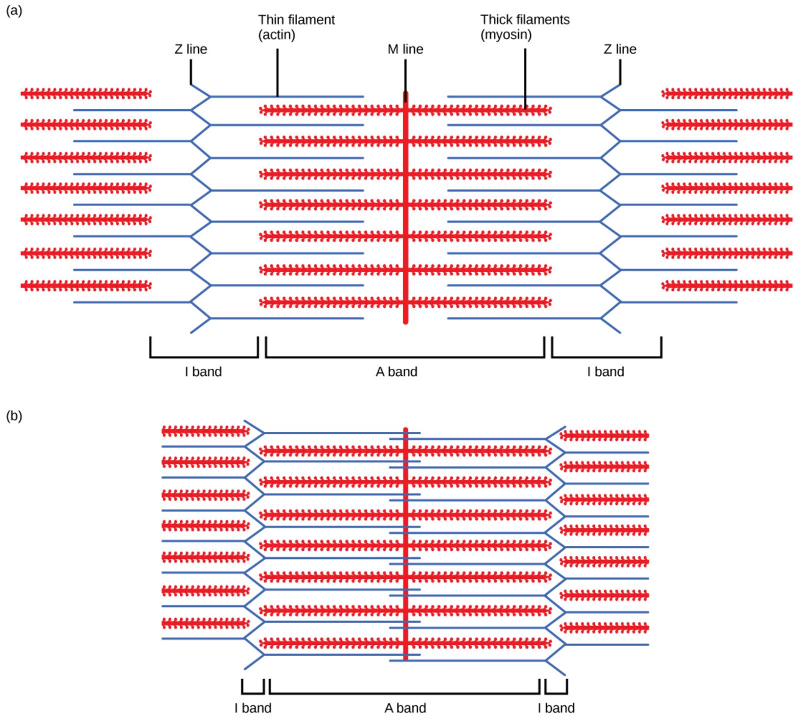
View this animation showing the organization of muscle fibers.

https://www.openstax.org/l/skeletal_muscle

When (a) a sarcomere (b) contracts, the Z lines move closer together and the I band gets smaller. The A band stays the same width and, at full contraction, the thin filaments overlap.

Sliding Filament Model of Contraction

For a muscle cell to contract, the sarcomere must shorten. However, thick and thin filaments—the components of sarcomeres—do not shorten. Instead, they slide by one another, causing the sarcomere to shorten while the filaments remain the same length. The sliding filament theory of muscle contraction was developed to fit the differences observed in the named bands on the sarcomere at different degrees of muscle contraction and relaxation. The mechanism of contraction is the binding of myosin to actin, forming cross-bridges that generate filament movement ([link]).



When a sarcomere shortens, some regions shorten whereas others stay the same length. A sarcomere is defined as the distance between two consecutive Z discs or Z lines; when a muscle contracts, the distance between the Z discs is reduced. The H zone—the central region of the A zone—contains only thick filaments and is shortened during contraction. The I band contains only thin filaments and also shortens. The A band does not shorten—it remains the same length—but A bands of different sarcomeres move closer together during contraction, eventually disappearing. Thin filaments are pulled by the thick filaments toward the center of the sarcomere until the Z discs approach the thick

filaments. The zone of overlap, in which thin filaments and thick filaments occupy the same area, increases as the thin filaments move inward.

ATP and Muscle Contraction

The motion of muscle shortening occurs as myosin heads bind to actin and pull the actin inwards. This action requires energy, which is provided by ATP. Myosin binds to actin at a binding site on the globular actin protein. Myosin has another binding site for ATP at which enzymatic activity hydrolyzes ATP to ADP, releasing an inorganic phosphate molecule and energy.

ATP binding causes myosin to release actin, allowing actin and myosin to detach from each other. After this happens, the newly bound ATP is converted to ADP and inorganic phosphate, P_i. The enzyme at the binding site on myosin is called ATPase. The energy released during ATP hydrolysis changes the angle of the myosin head into a "cocked" position. The myosin head is then in a position for further movement, possessing potential energy, but ADP and P_i are still attached. If actin binding sites are covered and unavailable, the myosin will remain in the high energy configuration with ATP hydrolyzed, but still attached.

If the actin binding sites are uncovered, a cross-

bridge will form; that is, the myosin head spans the distance between the actin and myosin molecules. P_i is then released, allowing myosin to expend the stored energy as a conformational change. The myosin head moves toward the M line, pulling the actin along with it. As the actin is pulled, the filaments move approximately 10 nm toward the M line. This movement is called the power stroke, as it is the step at which force is produced. As the actin is pulled toward the M line, the sarcomere shortens and the muscle contracts.

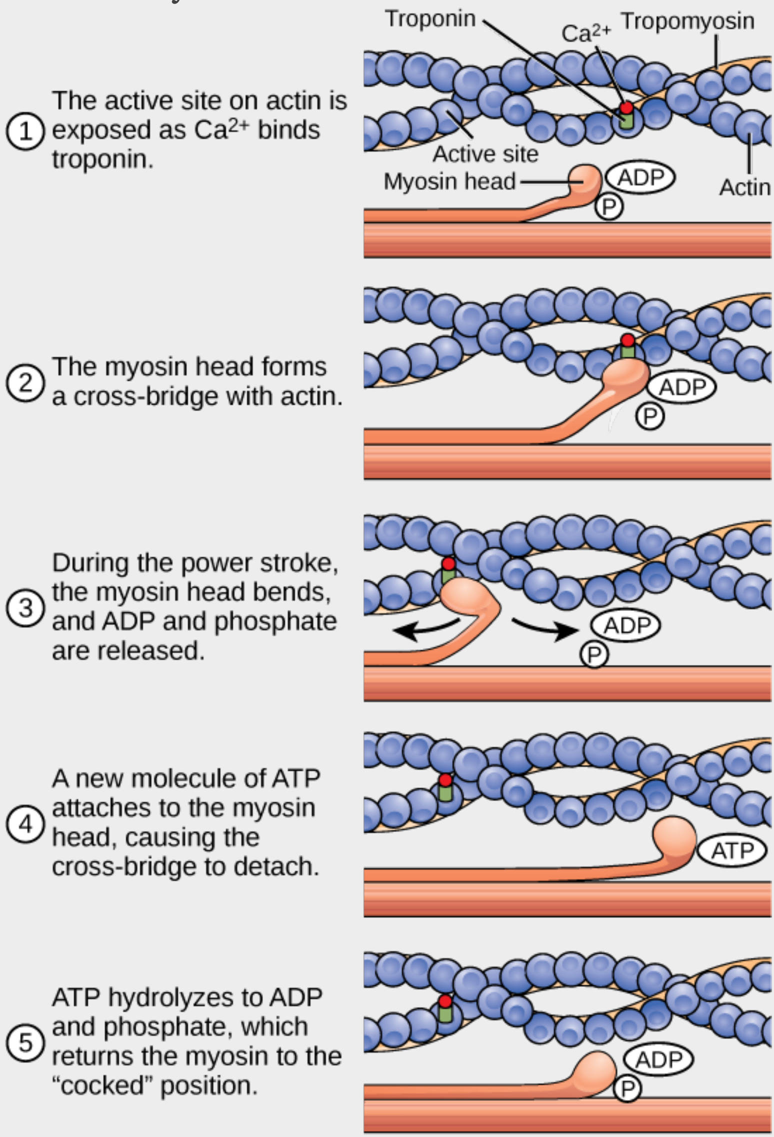
When the myosin head is “cocked,” it contains energy and is in a high-energy configuration. This energy is expended as the myosin head moves through the power stroke; at the end of the power stroke, the myosin head is in a low-energy position. After the power stroke, ADP is released; however, the cross-bridge formed is still in place, and actin and myosin are bound together. ATP can then attach to myosin, which allows the cross-bridge cycle to start again and further muscle contraction can occur ([link]).

Link to Learning

Watch this video explaining how a muscle contraction is signaled.
https://www.openstax.org/l/contract_muscle

Visual Connection

The cross-bridge muscle contraction cycle, which is triggered by Ca^{2+} binding to the actin active site, is shown. With each contraction cycle, actin moves relative to myosin.



Which of the following statements about muscle

contraction is true?

1. The power stroke occurs when ATP is hydrolyzed to ADP and phosphate.
2. The power stroke occurs when ADP and phosphate dissociate from the myosin head.
3. The power stroke occurs when ADP and phosphate dissociate from the actin active site.
4. The power stroke occurs when Ca^{2+} binds the calcium head.

Link to Learning

View this animation of the cross-bridge muscle contraction.

Regulatory Proteins

When a muscle is in a resting state, actin and myosin are separated. To keep actin from binding to the active site on myosin, regulatory proteins block the molecular binding sites. **Tropomyosin** blocks myosin binding sites on actin molecules, preventing cross-bridge formation and preventing contraction in a muscle without nervous input. **Troponin** binds

to tropomyosin and helps to position it on the actin molecule; it also binds calcium ions.

To enable a muscle contraction, tropomyosin must change conformation, uncovering the myosin-binding site on an actin molecule and allowing cross-bridge formation. This can only happen in the presence of calcium, which is kept at extremely low concentrations in the sarcoplasm. If present, calcium ions bind to troponin, causing conformational changes in troponin that allow tropomyosin to move away from the myosin binding sites on actin. Once the tropomyosin is removed, a cross-bridge can form between actin and myosin, triggering contraction. Cross-bridge cycling continues until Ca^{2+} ions and ATP are no longer available and tropomyosin again covers the binding sites on actin.

Excitation–Contraction Coupling

Excitation–contraction coupling is the link (transduction) between the action potential generated in the sarcolemma and the start of a muscle contraction. The trigger for calcium release from the sarcoplasmic reticulum into the sarcoplasm is a neural signal. Each skeletal muscle fiber is controlled by a motor neuron, which conducts signals from the brain or spinal cord to the muscle. The area of the sarcolemma on the muscle fiber that interacts with the neuron is called the **motor end**

plate. The end of the neuron's axon is called the synaptic terminal, and it does not actually contact the motor end plate. A small space called the synaptic cleft separates the synaptic terminal from the motor end plate. Electrical signals travel along the neuron's axon, which branches through the muscle and connects to individual muscle fibers at a neuromuscular junction.

The ability of cells to communicate electrically requires that the cells expend energy to create an electrical gradient across their cell membranes. This charge gradient is carried by ions, which are differentially distributed across the membrane. Each ion exerts an electrical influence and a concentration influence. Just as milk will eventually mix with coffee without the need to stir, ions also distribute themselves evenly, if they are permitted to do so. In this case, they are not permitted to return to an evenly mixed state.

The sodium–potassium ATPase uses cellular energy to move $K+$ ions inside the cell and $Na+$ ions outside. This alone accumulates a small electrical charge, but a big concentration gradient. There is lots of $K+$ in the cell and lots of $Na+$ outside the cell. Potassium is able to leave the cell through $K+$ channels that are open 90% of the time, and it does. However, $Na+$ channels are rarely open, so $Na+$ remains outside the cell. When $K+$ leaves the cell, obeying its concentration gradient, that effectively

leaves a negative charge behind. So at rest, there is a large concentration gradient for Na^+ to enter the cell, and there is an accumulation of negative charges left behind in the cell. This is the resting membrane potential. Potential in this context means a separation of electrical charge that is capable of doing work. It is measured in volts, just like a battery. However, the transmembrane potential is considerably smaller (0.07 V); therefore, the small value is expressed as millivolts (mV) or 70 mV. Because the inside of a cell is negative compared with the outside, a minus sign signifies the excess of negative charges inside the cell, −70 mV.

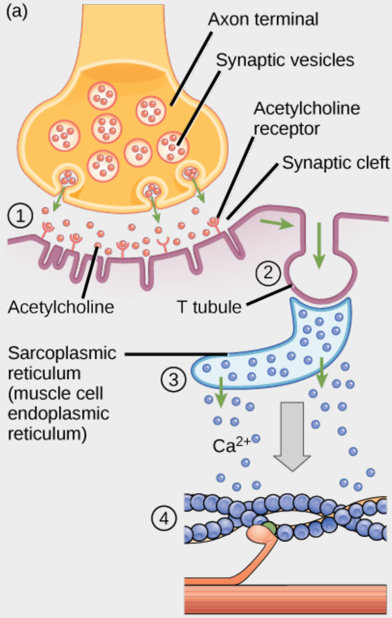
If an event changes the permeability of the membrane to Na^+ ions, they will enter the cell. That will change the voltage. This is an electrical event, called an action potential, that can be used as a cellular signal. Communication occurs between nerves and muscles through neurotransmitters. Neuron action potentials cause the release of neurotransmitters from the synaptic terminal into the synaptic cleft, where they can then diffuse across the synaptic cleft and bind to a receptor molecule on the motor end plate. The motor end plate possesses junctional folds—folds in the sarcolemma that create a large surface area for the neurotransmitter to bind to receptors. The receptors are actually sodium channels that open to allow the passage of Na^+ into the cell when they receive a neurotransmitter signal.

Acetylcholine (ACh) is a neurotransmitter released by motor neurons that binds to receptors in the motor end plate. Neurotransmitter release occurs when an action potential travels down the motor neuron's axon, resulting in altered permeability of the synaptic terminal membrane and an influx of calcium. The Ca^{2+} ions allow synaptic vesicles to move to and bind with the presynaptic membrane (on the neuron), and release neurotransmitter from the vesicles into the synaptic cleft. Once released by the synaptic terminal, ACh diffuses across the synaptic cleft to the motor end plate, where it binds with ACh receptors. As a neurotransmitter binds, these ion channels open, and Na^{+} ions cross the membrane into the muscle cell. This reduces the voltage difference between the inside and outside of the cell, which is called depolarization. As ACh binds at the motor end plate, this depolarization is called an end-plate potential. The depolarization then spreads along the sarcolemma, creating an action potential as sodium channels adjacent to the initial depolarization site sense the change in voltage and open. The action potential moves across the entire cell, creating a wave of depolarization.

ACh is broken down by the enzyme **acetylcholinesterase** (AChE) into acetyl and choline. AChE resides in the synaptic cleft, breaking down ACh so that it does not remain bound to ACh receptors, which would cause unwanted extended muscle contraction ([link]).

Visual Connection

This diagram shows excitation-contraction coupling in a skeletal muscle contraction. The sarcoplasmic reticulum is a specialized endoplasmic reticulum found in muscle cells.



1. Acetylcholine released from the axon terminal binds to receptors on the sarcolemma.
2. An action potential is generated and travels down the T tubule.
3. Ca^{2+} is released from the sarcoplasmic reticulum in response to the change in voltage.
4. Ca^{2+} binds troponin; Cross-bridges form between actin and myosin.
5. Acetylcholinesterase removes acetylcholine from the synaptic cleft.
6. Ca^{2+} is transported back into the sarcoplasmic reticulum.
7. Tropomyosin binds active sites on actin causing the cross-bridge to detach.

The deadly nerve gas Sarin irreversibly inhibits Acetylcholinesterase. What effect would Sarin have on muscle contraction?

After depolarization, the membrane returns to its

resting state. This is called repolarization, during which voltage-gated sodium channels close. Potassium channels continue at 90% conductance. Because the plasma membrane sodium–potassium ATPase always transports ions, the resting state (negatively charged inside relative to the outside) is restored. The period immediately following the transmission of an impulse in a nerve or muscle, in which a neuron or muscle cell regains its ability to transmit another impulse, is called the refractory period. During the refractory period, the membrane cannot generate another action potential. The refractory period allows the voltage-sensitive ion channels to return to their resting configurations. The sodium potassium ATPase continually moves Na^+ back out of the cell and K^+ back into the cell, and the K^+ leaks out leaving negative charge behind. Very quickly, the membrane repolarizes, so that it can again be depolarized.

Control of Muscle Tension

Neural control initiates the formation of actin–myosin cross-bridges, leading to the sarcomere shortening involved in muscle contraction. These contractions extend from the muscle fiber through connective tissue to pull on bones, causing skeletal movement. The pull exerted by a muscle is called tension, and the amount of force created by this tension can vary. This enables the same muscles to

move very light objects and very heavy objects. In individual muscle fibers, the amount of tension produced depends on the cross-sectional area of the muscle fiber and the frequency of neural stimulation.

The number of cross-bridges formed between actin and myosin determine the amount of tension that a muscle fiber can produce. Cross-bridges can only form where thick and thin filaments overlap, allowing myosin to bind to actin. If more cross-bridges are formed, more myosin will pull on actin, and more tension will be produced.

The ideal length of a sarcomere during production of maximal tension occurs when thick and thin filaments overlap to the greatest degree. If a sarcomere at rest is stretched past an ideal resting length, thick and thin filaments do not overlap to the greatest degree, and fewer cross-bridges can form. This results in fewer myosin heads pulling on actin, and less tension is produced. As a sarcomere is shortened, the zone of overlap is reduced as the thin filaments reach the H zone, which is composed of myosin tails. Because it is myosin heads that form cross-bridges, actin will not bind to myosin in this zone, reducing the tension produced by this myofiber. If the sarcomere is shortened even more, thin filaments begin to overlap with each other—reducing cross-bridge formation even further, and producing even less tension. Conversely, if the

sarcomere is stretched to the point at which thick and thin filaments do not overlap at all, no cross-bridges are formed and no tension is produced. This amount of stretching does not usually occur because accessory proteins, internal sensory nerves, and connective tissue oppose extreme stretching.

The primary variable determining force production is the number of myofibers within the muscle that receive an action potential from the neuron that controls that fiber. When using the biceps to pick up a pencil, the motor cortex of the brain only signals a few neurons of the biceps, and only a few myofibers respond. In vertebrates, each myofiber responds fully if stimulated. When picking up a piano, the motor cortex signals all of the neurons in the biceps and every myofiber participates. This is close to the maximum force the muscle can produce. As mentioned above, increasing the frequency of action potentials (the number of signals per second) can increase the force a bit more, because the tropomyosin is flooded with calcium.

Section Summary

The body contains three types of muscle tissue: skeletal muscle, cardiac muscle, and smooth muscle. Skeleton muscle tissue is composed of sarcomeres, the functional units of muscle tissue. Muscle contraction occurs when sarcomeres shorten, as

thick and thin filaments slide past each other, which is called the sliding filament model of muscle contraction. ATP provides the energy for cross-bridge formation and filament sliding. Regulatory proteins, such as troponin and tropomyosin, control cross-bridge formation. Excitation–contraction coupling transduces the electrical signal of the neuron, via acetylcholine, to an electrical signal on the muscle membrane, which initiates force production. The number of muscle fibers contracting determines how much force the whole muscle produces.

Visual Connection Questions

[link] Which of the following statements about muscle contraction is true?

1. The power stroke occurs when ATP is hydrolyzed to ADP and phosphate.
2. The power stroke occurs when ADP and phosphate dissociate from the myosin head.
3. The power stroke occurs when ADP and phosphate dissociate from the actin active site.
4. The power stroke occurs when Ca^{2+} binds the calcium head.

[link] B

[link] The deadly nerve gas Sarin irreversibly inhibits Acetylcholinesterase. What effect would Sarin have on muscle contraction?

[link] In the presence of Sarin, acetycholine is not removed from the synapse, resulting in continuous stimulation of the muscle plasma membrane. At first, muscle activity is intense and uncontrolled, but the ion gradients dissipate, so electrical signals in the T-tubules are no longer possible. The result is paralysis, leading to death by asphyxiation.

Review Questions

In relaxed muscle, the myosin-binding site on actin is blocked by ______.

1. titin
2. troponin
3. myoglobin
4. tropomyosin

D

The cell membrane of a muscle fiber is called a ______.

1. myofibril
2. sarcolemma
3. sarcoplasm
4. myofilament

B

The muscle relaxes if no new nerve signal arrives. However the neurotransmitter from the previous stimulation is still present in the synapse. The activity of ______ helps to remove this neurotransmitter.

1. myosin
2. action potential
3. tropomyosin
4. acetylcholinesterase

D

The ability of a muscle to generate tension immediately after stimulation is dependent on:

1. myosin interaction with the M line
2. overlap of myosin and actin
3. actin attachments to the Z line
4. none of the above

D

Botulinum toxin causes flaccid paralysis of the muscles, and is used for cosmetic purposes under the name Botox. Which of the following is the most likely mechanism of action of Botox?

1. Botox decreases the production of acetylcholinesterase.
2. Botox increases calcium release from the sarcoplasmic reticulum.
3. Botox blocks the ATP binding site in actin.
4. Botox decreases the release of acetylcholine from motor neurons.

D

Critical Thinking Questions

How would muscle contractions be affected if ATP was completely depleted in a muscle fiber?

Because ATP is required for myosin to release from actin, muscles would remain rigidly contracted until more ATP was available for the myosin cross-bridge release. This is why dead vertebrates undergo rigor mortis.

What factors contribute to the amount of tension produced in an individual muscle fiber?

The cross-sectional area, the length of the muscle fiber at rest, and the frequency of neural stimulation.

What effect will low blood calcium have on neurons? What effect will low blood calcium have on skeletal muscles?

Neurons will not be able to release neurotransmitter without calcium. Skeletal muscles have calcium stored and don’t need any from the outside.

Skeletal muscles can only produce a mechanical

force as they are contracted, but a leg flexes and extends while walking. How can muscles perform this task?

Muscles are able to drive locomotion (and other tasks involving opposing motions) because they are paired. When walking, the hamstring muscle contracts first, causing the leg to flex around the knee joint. The quadriceps muscle then contracts (while the hamstring relaxes and extends) to straighten the leg as the foot returns to the ground.

Glossary

actin
: globular contractile protein that interacts with myosin for muscle contraction

acetylcholinesterase
: (AChE) enzyme that breaks down ACh into acetyl and choline

cardiac muscle tissue
: muscle tissue found only in the heart; cardiac contractions pump blood throughout the body and maintain blood pressure

motor end plate
: sarcolemma of the muscle fiber that interacts

with the neuron

myofibril
long cylindrical structures that lie parallel to the muscle fiber

myofilament
small structures that make up myofibrils

myosin
contractile protein that interacts with actin for muscle contraction

sarcolemma
plasma membrane of a skeletal muscle fiber

sarcomere
functional unit of skeletal muscle

skeletal muscle tissue
forms skeletal muscles, which attach to bones and control locomotion and any movement that can be consciously controlled

smooth muscle tissue
occurs in the walls of hollow organs such as the intestines, stomach, and urinary bladder, and around passages such as the respiratory tract and blood vessels

thick filament
a group of myosin molecules

thin filament
: two polymers of actin wound together along with tropomyosin and troponin

tropomyosin
: acts to block myosin binding sites on actin molecules, preventing cross-bridge formation and preventing contraction until a muscle receives a neuron signal

troponin
: binds to tropomyosin and helps to position it on the actin molecule, and also binds calcium ions

Introduction

class="introduction" Lungs, which appear as nearly transparent tissue surrounding the heart in this X-ray of a dog (left), are the central organs of the respiratory system. The left lung is smaller than the right lung to accommodate space for the heart. A dog’s nose (right) has a slit on the side of each nostril. When tracking a scent, the slits open, blocking the front of the nostrils. This allows the dog to exhale though the now-open area on the side of the nostrils without losing the scent that is being followed. (credit a: modification of work by Geoff Stearns; credit b: modification of work by Cory Zanker)

Breathing is an involuntary event. How often a breath is taken and how much air is inhaled or exhaled are tightly regulated by the respiratory center in the brain. Humans, when they aren’t exerting themselves, breathe approximately 15 times per minute on average. Canines, like the dog in [link], have a respiratory rate of about 15–30

breaths per minute. With every inhalation, air fills the lungs, and with every exhalation, air rushes back out. That air is doing more than just inflating and deflating the lungs in the chest cavity. The air contains oxygen that crosses the lung tissue, enters the bloodstream, and travels to organs and tissues. Oxygen (O_2) enters the cells where it is used for metabolic reactions that produce ATP, a high-energy compound. At the same time, these reactions release carbon dioxide (CO_2) as a by-product. CO_2 is toxic and must be eliminated. Carbon dioxide exits the cells, enters the bloodstream, travels back to the lungs, and is expired out of the body during exhalation.

Systems of Gas Exchange

By the end of this section, you will be able to do the following:

- Describe the passage of air from the outside environment to the lungs
- Explain how the lungs are protected from particulate matter

The primary function of the respiratory system is to deliver oxygen to the cells of the body's tissues and remove carbon dioxide, a cell waste product. The main structures of the human respiratory system are the nasal cavity, the trachea, and lungs.

All aerobic organisms require oxygen to carry out their metabolic functions. Along the evolutionary tree, different organisms have devised different means of obtaining oxygen from the surrounding atmosphere. The environment in which the animal lives greatly determines how an animal respires. The complexity of the respiratory system is correlated with the size of the organism. As animal size increases, diffusion distances increase and the ratio of surface area to volume drops. In unicellular organisms, diffusion across the cell membrane is sufficient for supplying oxygen to the cell ([link]). Diffusion is a slow, passive transport process. In order for diffusion to be a feasible means of providing oxygen to the cell, the rate of oxygen uptake must match the rate of diffusion across the

membrane. In other words, if the cell were very large or thick, diffusion would not be able to provide oxygen quickly enough to the inside of the cell. Therefore, dependence on diffusion as a means of obtaining oxygen and removing carbon dioxide remains feasible only for small organisms or those with highly-flattened bodies, such as many flatworms (Platyhelminthes). Larger organisms had to evolve specialized respiratory tissues, such as gills, lungs, and respiratory passages accompanied by complex circulatory systems, to transport oxygen throughout their entire body.

The cell of the unicellular alga *Ventricaria ventricosa* is one of the largest known, reaching one to five centimeters in diameter. Like all single-celled organisms, *V. ventricosa* exchanges gases across the cell membrane.

This flatworm's process of respiration works by

diffusion across the outer membrane. (credit: Stephen Childs)

Direct Diffusion

For small multicellular organisms, diffusion across the outer membrane is sufficient to meet their oxygen needs. Gas exchange by direct diffusion across surface membranes is efficient for organisms less than 1 mm in diameter. In simple organisms, such as cnidarians and flatworms, every cell in the body is close to the external environment. Their cells are kept moist and gases diffuse quickly via direct diffusion. Flatworms are small, literally flat worms, which 'breathe' through diffusion across the outer membrane ([link]). The flat shape of these organisms increases the surface area for diffusion, ensuring that each cell within the body is close to the outer membrane surface and has access to oxygen. If the flatworm had a cylindrical body, then the cells in the center would not be able to get oxygen.

This common carp, like many other aquatic organisms, has gills that allow it to obtain oxygen from water. (credit: "Guitardude012"/Wikimedia Commons) As water flows over the gills, oxygen is transferred to blood via the veins. (credit "fish": modification of work by Duane Raver, NOAA)

Skin and Gills

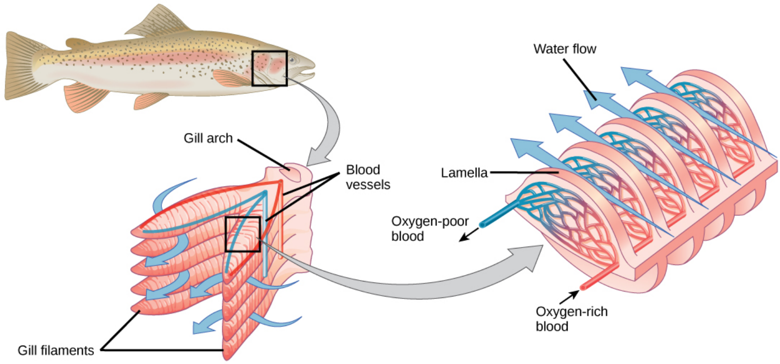
Earthworms and amphibians use their skin (integument) as a respiratory organ. A dense network of capillaries lies just below the skin and facilitates gas exchange between the external environment and the circulatory system. The respiratory surface must be kept moist in order for the gases to dissolve and diffuse across cell membranes.

Organisms that live in water need to obtain oxygen from the water. Oxygen dissolves in water but at a

lower concentration than in the atmosphere. The atmosphere has roughly 21 percent oxygen. In water, the oxygen concentration is much lower than that. Fish and many other aquatic organisms have evolved gills to take up the dissolved oxygen from water ([link]). Gills are thin tissue filaments that are highly branched and folded. When water passes over the gills, the dissolved oxygen in water rapidly diffuses across the gills into the bloodstream. The circulatory system can then carry the oxygenated blood to the other parts of the body. In animals that contain coelomic fluid instead of blood, oxygen diffuses across the gill surfaces into the coelomic fluid. Gills are found in mollusks, annelids, and crustaceans.

The folded surfaces of the gills provide a large surface area to ensure that the fish gets sufficient

oxygen. Diffusion is a process in which material travels from regions of high concentration to low concentration until equilibrium is reached. In this case, blood with a low concentration of oxygen molecules circulates through the gills. The concentration of oxygen molecules in water is higher than the concentration of oxygen molecules in gills. As a result, oxygen molecules diffuse from water (high concentration) to blood (low concentration), as shown in [link]. Similarly, carbon dioxide molecules in the blood diffuse from the blood (high concentration) to water (low concentration).



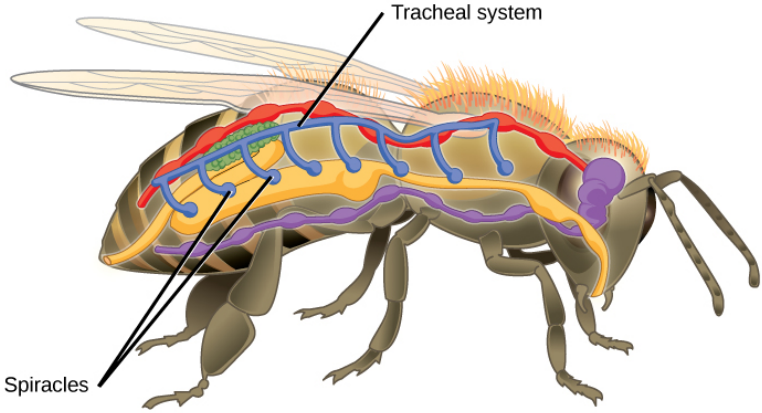
Insects perform respiration via a tracheal system.

Tracheal Systems

Insect respiration is independent of its circulatory system; therefore, the blood does not play a direct role in oxygen transport. Insects have a highly specialized type of respiratory system called the tracheal system, which consists of a network of small tubes that carries oxygen to the entire body.

The tracheal system is the most direct and efficient respiratory system in active animals. The tubes in the tracheal system are made of a polymeric material called chitin.

Insect bodies have openings, called spiracles, along the thorax and abdomen. These openings connect to the tubular network, allowing oxygen to pass into the body ([link]) and regulating the diffusion of CO_2 and water vapor. Air enters and leaves the tracheal system through the spiracles. Some insects can ventilate the tracheal system with body movements.



The trachea and bronchi are made of incomplete rings of cartilage. (credit: modification of work by Gray's Anatomy) The trachea bifurcates into the right and left bronchi in the lungs. The right lung is made of three lobes and is larger. To accommodate the heart, the left lung is smaller and has only two lobes. Terminal bronchioles are connected by respiratory bronchioles to alveolar ducts and alveolar sacs. Each alveolar sac contains 20 to 30 spherical alveoli and has the appearance of a bunch

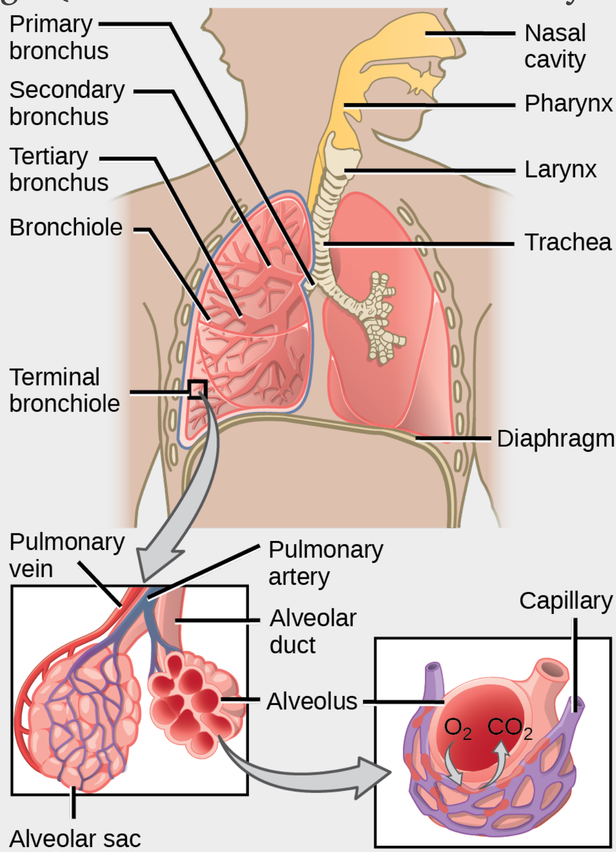
of grapes. Air flows into the atrium of the alveolar sac, then circulates into alveoli where gas exchange occurs with the capillaries. Mucous glands secrete mucous into the airways, keeping them moist and flexible. (credit: modification of work by Mariana Ruiz Villareal)

Mammalian Systems

In mammals, pulmonary ventilation occurs via inhalation (breathing). During inhalation, air enters the body through the **nasal cavity** located just inside the nose ([link]). As air passes through the nasal cavity, the air is warmed to body temperature and humidified. The respiratory tract is coated with mucus to seal the tissues from direct contact with air. Mucus is high in water. As air crosses these surfaces of the mucous membranes, it picks up water. These processes help equilibrate the air to the body conditions, reducing any damage that cold, dry air can cause. Particulate matter that is floating in the air is removed in the nasal passages via mucus and cilia. The processes of warming, humidifying, and removing particles are important protective mechanisms that prevent damage to the trachea and lungs. Thus, inhalation serves several purposes in addition to bringing oxygen into the respiratory system.

Visual Connection

Air enters the respiratory system through the nasal cavity and pharynx, and then passes through the trachea and into the bronchi, which bring air into the lungs. (credit: modification of work by NCI)

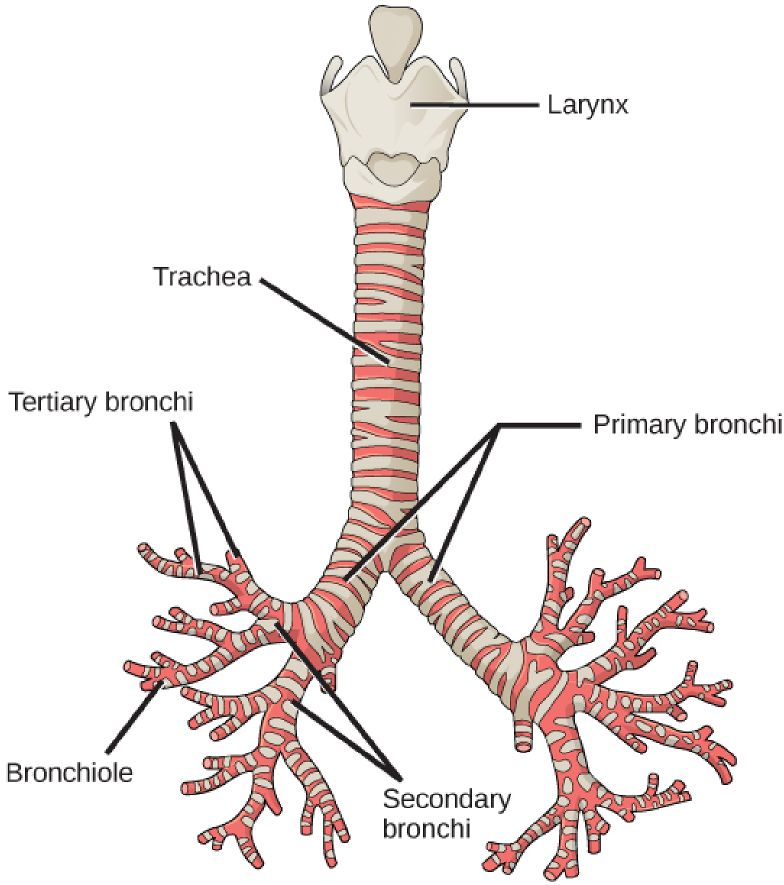


Which of the following statements about the mammalian respiratory system is false?

1. When we breathe in, air travels from the pharynx to the trachea.
2. The bronchioles branch into bronchi.
3. Alveolar ducts connect to alveolar sacs.
4. Gas exchange between the lung and blood

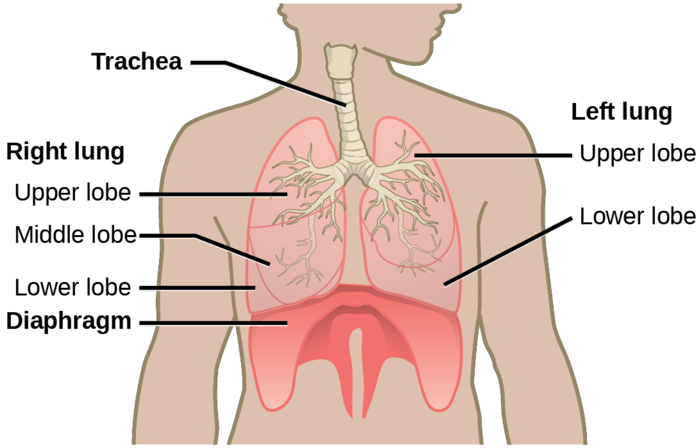
takes place in the alveolus.

From the nasal cavity, air passes through the **pharynx** (throat) and the **larynx** (voice box), as it makes its way to the **trachea** ([link]). The main function of the trachea is to funnel the inhaled air to the lungs and the exhaled air back out of the body. The human trachea is a cylinder about 10 to 12 cm long and 2 cm in diameter that sits in front of the esophagus and extends from the larynx into the chest cavity where it divides into the two primary bronchi at the midthorax. It is made of incomplete rings of hyaline cartilage and smooth muscle ([link]). The trachea is lined with mucus-producing goblet cells and ciliated epithelia. The cilia propel foreign particles trapped in the mucus toward the pharynx. The cartilage provides strength and support to the trachea to keep the passage open. The smooth muscle can contract, decreasing the trachea’s diameter, which causes expired air to rush upwards from the lungs at a great force. The forced exhalation helps expel mucus when we cough. Smooth muscle can contract or relax, depending on stimuli from the external environment or the body’s nervous system.



Lungs: Bronchi and Alveoli

The end of the trachea bifurcates (divides) to the right and left lungs. The lungs are not identical. The right lung is larger and contains three lobes, whereas the smaller left lung contains two lobes ([link]). The muscular **diaphragm**, which facilitates breathing, is inferior to (below) the lungs and marks the end of the thoracic cavity.

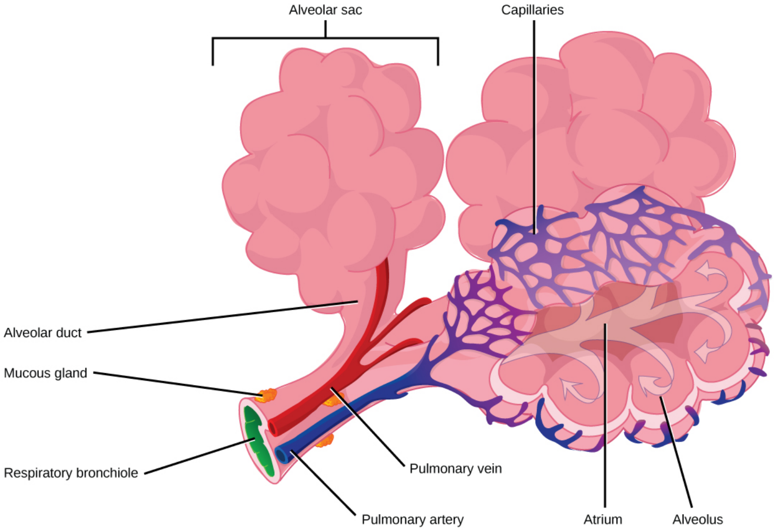


In the lungs, air is diverted into smaller and smaller passages, or **bronchi**. Air enters the lungs through the two **primary (main) bronchi** (singular: bronchus). Each bronchus divides into secondary bronchi, then into tertiary bronchi, which in turn divide, creating smaller and smaller diameter **bronchioles** as they split and spread through the lung. Like the trachea, the bronchi are made of cartilage and smooth muscle. At the bronchioles, the cartilage is replaced with elastic fibers. Bronchi are innervated by nerves of both the parasympathetic and sympathetic nervous systems that control muscle contraction (parasympathetic) or relaxation (sympathetic) in the bronchi and bronchioles, depending on the nervous system's cues. In humans, bronchioles with a diameter smaller than 0.5 mm are the **respiratory bronchioles**. They lack cartilage and therefore rely on inhaled air to support their shape. As the passageways decrease in diameter, the relative amount of smooth muscle

increases.

The **terminal bronchioles** subdivide into microscopic branches called respiratory bronchioles. The respiratory bronchioles subdivide into several alveolar ducts. Numerous alveoli and alveolar sacs surround the alveolar ducts. The alveolar sacs resemble bunches of grapes tethered to the end of the bronchioles ([link]). In the acinar region, the **alveolar ducts** are attached to the end of each bronchiole. At the end of each duct are approximately 100 **alveolar sacs**, each containing 20 to 30 **alveoli** that are 200 to 300 microns in diameter. Gas exchange occurs only in alveoli. Alveoli are made of thin-walled parenchymal cells, typically one-cell thick, that look like tiny bubbles within the sacs. Alveoli are in direct contact with capillaries (one-cell thick) of the circulatory system. Such intimate contact ensures that oxygen will diffuse from alveoli into the blood and be distributed to the cells of the body. In addition, the carbon dioxide that was produced by cells as a waste product will diffuse from the blood into alveoli to be exhaled. The anatomical arrangement of capillaries and alveoli emphasizes the structural and functional relationship of the respiratory and circulatory systems. Because there are so many alveoli (~300 million per lung) within each alveolar sac and so many sacs at the end of each alveolar duct, the lungs have a sponge-like consistency. This organization produces a very large surface area that

is available for gas exchange. The surface area of alveoli in the lungs is approximately 75 m^2. This large surface area, combined with the thin-walled nature of the alveolar parenchymal cells, allows gases to easily diffuse across the cells.



Link to Learning

Watch the following video to review the respiratory system.
https://www.openstax.org/l/lungs_pulmonary

The bronchi and bronchioles contain cilia that help move mucus and other particles out of the lungs. (credit: Louisa Howard, modification of work by Dartmouth Electron Microscope Facility)

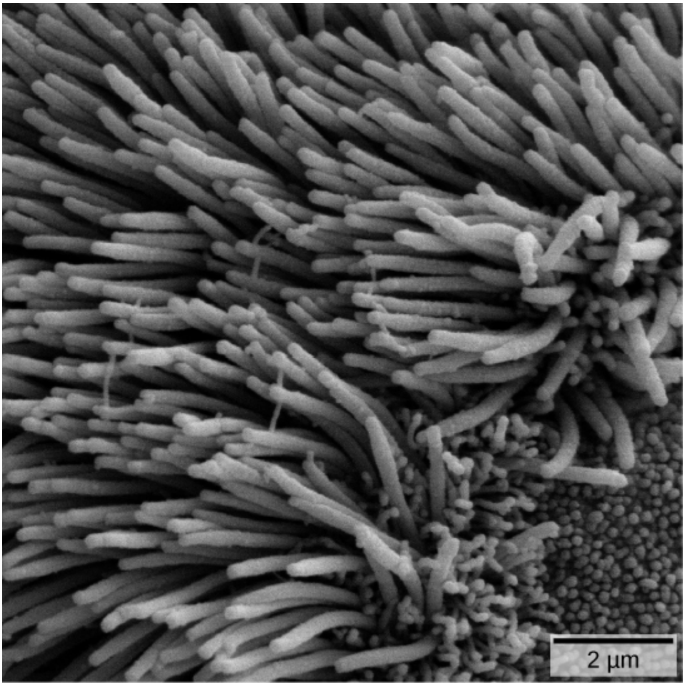
Protective Mechanisms

The air that organisms breathe contains **particulate matter** such as dust, dirt, viral particles, and bacteria that can damage the lungs or trigger allergic immune responses. The respiratory system contains several protective mechanisms to avoid problems or tissue damage. In the nasal cavity, hairs and mucus trap small particles, viruses, bacteria, dust, and dirt to prevent their entry.

If particulates do make it beyond the nose, or enter through the mouth, the bronchi and bronchioles of the lungs also contain several protective devices. The lungs produce **mucus**—a sticky substance made of **mucin**, a complex glycoprotein, as well as salts and water—that traps particulates. The bronchi and bronchioles contain cilia, small hair-like projections that line the walls of the bronchi and bronchioles ([link]). These cilia beat in unison and move mucus and particles out of the bronchi and bronchioles back up to the throat where it is swallowed and eliminated via the esophagus.

In humans, for example, tar and other substances in cigarette smoke destroy or paralyze the cilia, making the removal of particles more difficult. In addition, smoking causes the lungs to produce more mucus, which the damaged cilia are not able to move. This causes a persistent cough, as the lungs try to rid themselves of particulate matter, and

makes smokers more susceptible to respiratory ailments.



Section Summary

Animal respiratory systems are designed to facilitate gas exchange. In mammals, air is warmed and humidified in the nasal cavity. Air then travels down the pharynx, through the trachea, and into the lungs. In the lungs, air passes through the branching bronchi, reaching the respiratory bronchioles, which house the first site of gas exchange. The respiratory bronchioles open into the alveolar ducts, alveolar sacs, and alveoli. Because there are so many alveoli and alveolar sacs in the lung, the surface area for

gas exchange is very large. Several protective mechanisms are in place to prevent damage or infection. These include the hair and mucus in the nasal cavity that trap dust, dirt, and other particulate matter before they can enter the system. In the lungs, particles are trapped in a mucus layer and transported via cilia up to the esophageal opening at the top of the trachea to be swallowed.

[link] Which of the following statements about the mammalian respiratory system is false?

1. When we breathe in, air travels from the pharynx to the trachea.
2. The bronchioles branch into bronchi.
3. Alveolar ducts connect to alveolar sacs.
4. Gas exchange between the lung and blood takes place in the alveolus.

[link] B

Review Questions

The respiratory system _____.

1. provides body tissues with oxygen

2. provides body tissues with oxygen and carbon dioxide
3. establishes how many breaths are taken per minute
4. provides the body with carbon dioxide

A

Air is warmed and humidified in the nasal passages. This helps to _____.

1. ward off infection
2. decrease sensitivity during breathing
3. prevent damage to the lungs
4. all of the above

C

Which is the order of airflow during inhalation?

1. nasal cavity, trachea, larynx, bronchi, bronchioles, alveoli
2. nasal cavity, larynx, trachea, bronchi, bronchioles, alveoli
3. nasal cavity, larynx, trachea, bronchioles, bronchi, alveoli
4. nasal cavity, trachea, larynx, bronchioles, bronchi, alveoli

B

Critical Thinking Questions

Describe the function of these terms and describe where they are located: main bronchus, trachea, alveoli, and acinus.

The main bronchus is the conduit in the lung that funnels air to the airways where gas exchange occurs. The main bronchus attaches the lungs to the very end of the trachea where it bifurcates. The trachea is the cartilaginous structure that extends from the pharynx to the primary bronchi. It serves to funnel air to the lungs. The alveoli are the sites of gas exchange; they are located at the terminal regions of the lung and are attached to the respiratory bronchioles. The acinus is the structure in the lung where gas exchange occurs.

How does the structure of alveoli maximize gas exchange?

The sac-like structure of the alveoli increases

their surface area. In addition, the alveoli are made of thin-walled parenchymal cells. These features allow gases to easily diffuse across the cells.

Glossary

alveolar duct
: duct that extends from the terminal bronchiole to the alveolar sac

alveolar sac
: structure consisting of two or more alveoli that share a common opening

alveolus
: (plural: alveoli) (also, air sac) terminal region of the lung where gas exchange occurs

bronchus
: (plural: bronchi) smaller branch of cartilaginous tissue that stems off of the trachea; air is funneled through the bronchi to the region where gas exchange occurs in alveoli

bronchiole
: airway that extends from the main tertiary bronchi to the alveolar sac

diaphragm

domed-shaped skeletal muscle located under lungs that separates the thoracic cavity from the abdominal cavity

larynx
: voice box, a short passageway connecting the pharynx and the trachea

mucin
: complex glycoprotein found in mucus

mucus
: sticky protein-containing fluid secretion in the lung that traps particulate matter to be expelled from the body

nasal cavity
: opening of the respiratory system to the outside environment

particulate matter
: small particle such as dust, dirt, viral particles, and bacteria that are in the air

pharynx
: throat; a tube that starts in the internal nares and runs partway down the neck, where it opens into the esophagus and the larynx

primary bronchus
: (also, main bronchus) region of the airway within the lung that attaches to the trachea

and bifurcates to each lung where it branches into secondary bronchi

respiratory bronchiole
: terminal portion of the bronchiole tree that is attached to the terminal bronchioles and alveoli ducts, alveolar sacs, and alveoli

terminal bronchiole
: region of bronchiole that attaches to the respiratory bronchioles

trachea
: cartilaginous tube that transports air from the larynx to the primary bronchi

Gas Exchange across Respiratory Surfaces

By the end of this section, you will be able to do the following:

- Name and describe lung volumes and capacities
- Understand how gas pressure influences how gases move into and out of the body

The structure of the lung maximizes its surface area to increase gas diffusion. Because of the enormous number of alveoli (approximately 300 million in each human lung), the surface area of the lung is very large (75 m^2). Having such a large surface area increases the amount of gas that can diffuse into and out of the lungs.

Basic Principles of Gas Exchange

Gas exchange during respiration occurs primarily through diffusion. Diffusion is a process in which transport is driven by a concentration gradient. Gas molecules move from a region of high concentration to a region of low concentration. Blood that is low in oxygen concentration and high in carbon dioxide concentration undergoes gas exchange with air in the lungs. The air in the lungs has a higher concentration of oxygen than that of oxygen-depleted blood and a lower concentration of carbon dioxide. This concentration gradient allows for gas

exchange during respiration.

Partial pressure is a measure of the concentration of the individual components in a mixture of gases. The total pressure exerted by the mixture is the sum of the partial pressures of the components in the mixture. The rate of diffusion of a gas is proportional to its partial pressure within the total gas mixture. This concept is discussed further in detail below.

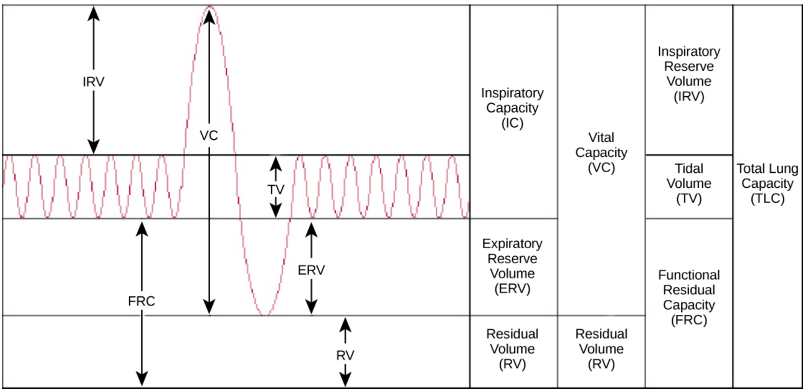
Human lung volumes and capacities are shown. The total lung capacity of the adult male is six liters. Tidal volume is the volume of air inhaled in a single, normal breath. Inspiratory capacity is the amount of air taken in during a deep breath, and residual volume is the amount of air left in the lungs after forceful respiration.

Lung Volumes and Capacities

Different animals have different lung capacities based on their activities. Cheetahs have evolved a much higher lung capacity than humans; it helps provide oxygen to all the muscles in the body and allows them to run very fast. Elephants also have a high lung capacity. In this case, it is not because they run fast but because they have a large body and must be able to take up oxygen in accordance with their body size.

Human lung size is determined by genetics, sex, and

height. At maximal capacity, an average lung can hold almost six liters of air, but lungs do not usually operate at maximal capacity. Air in the lungs is measured in terms of **lung volumes** and **lung capacities** ([link] and [link]). Volume measures the amount of air for one function (such as inhalation or exhalation). Capacity is any two or more volumes (for example, how much can be inhaled from the end of a maximal exhalation).



Lung Volumes and Capacities (Avg Adult Male)			
Volume/ Capacity	**Definition**	**Volume (liters)**	**Equations**

Tidal volume (TV)	Amount of air inhaled during a normal breath	0.5	-	
Expiratory reserve volume (ERV)	Amount of air that can be exhaled after a normal exhalation	1.2	-	
Inspiratory reserve volume (IRV)	Amount of air that can be further inhaled after a normal inhalation	3.1	-	
Residual volume (RV)	Air left in the lungs after a forced exhalation	1.2	-	
Vital capacity (VC)	Maximum amount of air that can be moved in or out of the lungs in a single respiratory cycle	4.8	ERV + TV + IRV	

Inspiratory capacity (IC)	Volume of air that can be inhaled in addition to a normal exhalation	3.6	TV + IRV
Functional residual capacity (FRC)	Volume of air remaining after a normal exhalation	2.4	ERV + RV
Total lung capacity (TLC)	Total volume of air in the lungs after a maximal inspiration	6.0	RV + ERV + TV + IRV
Forced expiratory volume (FEV1)	How much air can be forced out of the lungs over a specific time period, usually one second	~4.1 to 5.5	-

The volume in the lung can be divided into four units: tidal volume, expiratory reserve volume, inspiratory reserve volume, and residual volume.

Tidal volume (TV) measures the amount of air that is inspired and expired during a normal breath. On average, this volume is around one-half liter, which is a little less than the capacity of a 20-ounce drink bottle. The **expiratory reserve volume (ERV)** is the additional amount of air that can be exhaled after a normal exhalation. It is the reserve amount that can be exhaled beyond what is normal. Conversely, the **inspiratory reserve volume (IRV)** is the additional amount of air that can be inhaled after a normal inhalation. The **residual volume (RV)** is the amount of air that is left after expiratory reserve volume is exhaled. The lungs are never completely empty: There is always some air left in the lungs after a maximal exhalation. If this residual volume did not exist and the lungs emptied completely, the lung tissues would stick together and the energy necessary to reinflate the lung could be too great to overcome. Therefore, there is always some air remaining in the lungs. Residual volume is also important for preventing large fluctuations in respiratory gases (O_2 and CO_2). The residual volume is the only lung volume that cannot be measured directly because it is impossible to completely empty the lung of air. This volume can only be calculated rather than measured.

Capacities are measurements of two or more volumes. The **vital capacity (VC)** measures the maximum amount of air that can be inhaled or exhaled during a respiratory cycle. It is the sum of

the expiratory reserve volume, tidal volume, and inspiratory reserve volume. The **inspiratory capacity (IC)** is the amount of air that can be inhaled after the end of a normal expiration. It is, therefore, the sum of the tidal volume and inspiratory reserve volume. The **functional residual capacity (FRC)** includes the expiratory reserve volume and the residual volume. The FRC measures the amount of additional air that can be exhaled after a normal exhalation. Lastly, the **total lung capacity (TLC)** is a measurement of the total amount of air that the lung can hold. It is the sum of the residual volume, expiratory reserve volume, tidal volume, and inspiratory reserve volume.

Lung volumes are measured by a technique called **spirometry**. An important measurement taken during spirometry is the **forced expiratory volume (FEV)**, which measures how much air can be forced out of the lung over a specific period, usually one second (FEV1). In addition, the forced vital capacity (FVC), which is the total amount of air that can be forcibly exhaled, is measured. The ratio of these values (**FEV1/FVC ratio**) is used to diagnose lung diseases including asthma, emphysema, and fibrosis. If the FEV1/FVC ratio is high, the lungs are not compliant (meaning they are stiff and unable to bend properly), and the patient most likely has lung fibrosis. Patients exhale most of the lung volume very quickly. Conversely, when the FEV1/FVC ratio is low, there is resistance in the lung that is

characteristic of asthma. In this instance, it is hard for the patient to get the air out of his or her lungs, and it takes a long time to reach the maximal exhalation volume. In either case, breathing is difficult and complications arise.

Career Connection

Respiratory Therapist

Respiratory therapists or respiratory practitioners evaluate and treat patients with lung and cardiovascular diseases. They work as part of a medical team to develop treatment plans for patients. Respiratory therapists may treat premature babies with underdeveloped lungs, patients with chronic conditions such as asthma, or older patients suffering from lung disease such as emphysema and chronic obstructive pulmonary disease (COPD). They may operate advanced equipment such as compressed gas delivery systems, ventilators, blood gas analyzers, and resuscitators. Specialized programs to become a respiratory therapist generally lead to a bachelor's degree with a respiratory therapist specialty. Because of a growing aging population, career opportunities as a respiratory therapist are expected to remain strong.

Gas Pressure and Respiration

The respiratory process can be better understood by examining the properties of gases. Gases move freely, but gas particles are constantly hitting the walls of their vessel, thereby producing gas pressure.

Air is a mixture of gases, primarily nitrogen (N_2; 78.6 percent), oxygen (O_2; 20.9 percent), water vapor (H_2O; 0.5 percent), and carbon dioxide (CO_2; 0.04 percent). Each gas component of that mixture exerts a pressure. The pressure for an individual gas in the mixture is the partial pressure of that gas. Approximately 21 percent of atmospheric gas is oxygen. Carbon dioxide, however, is found in relatively small amounts, 0.04 percent. The partial pressure for oxygen is much greater than that of carbon dioxide. The partial pressure of any gas can be calculated by:

$$P = (P_{atm}) \times (\text{percent content in mixture}).$$

P_{atm}, the atmospheric pressure, is the sum of all of the partial pressures of the atmospheric gases added together,

$$P_{atm} = P_{N_2} + P_{O_2} + P_{H_2O} + P_{CO_2} = 760 \text{ mm Hg}$$

$\times$ (percent content in mixture).

The pressure of the atmosphere at sea level is 760 mm Hg. Therefore, the partial pressure of oxygen is:
$$P_{O_2} = (760 \text{ mm Hg})(0.21) = 160 \text{ mm Hg}$$

and for carbon dioxide:
$$P_{CO_2} = (760 \text{ mm Hg})(0.0004) = 0.3 \text{ mm Hg}.$$

At high altitudes, P_{atm} decreases but concentration does not change; the partial pressure decrease is due to the reduction in P_{atm}.

When the air mixture reaches the lung, it has been humidified. The pressure of the water vapor in the lung does not change the pressure of the air, but it must be included in the partial pressure equation. For this calculation, the water pressure (47 mm Hg) is subtracted from the atmospheric pressure:
$$760 \text{ mm Hg} - 47 \text{ mm Hg} = 713 \text{ mm Hg}$$

and the partial pressure of oxygen is:
$$(760 \text{ mm Hg} - 47 \text{ mm Hg}) \times 0.21 = 150 \text{ mm Hg}.$$

These pressures determine the gas exchange, or the flow of gas, in the system. Oxygen and carbon dioxide will flow according to their pressure gradient from high to low. Therefore, understanding the partial pressure of each gas will aid in understanding how gases move in the respiratory system.

Gas Exchange across the Alveoli

In the body, oxygen is used by cells of the body's tissues and carbon dioxide is produced as a waste product. The ratio of carbon dioxide production to oxygen consumption is the **respiratory quotient (RQ)**. RQ varies between 0.7 and 1.0. If just glucose were used to fuel the body, the RQ would equal one. One mole of carbon dioxide would be produced for every mole of oxygen consumed. Glucose, however, is not the only fuel for the body. Protein and fat are also used as fuels for the body. Because of this, less carbon dioxide is produced than oxygen is consumed and the RQ is, on average, about 0.7 for fat and about 0.8 for protein.

The RQ is used to calculate the partial pressure of oxygen in the alveolar spaces within the lung, the **alveolar P_{O_2}**. Above, the partial pressure of oxygen in the lungs was calculated to be 150 mm Hg. However, lungs never fully deflate with an exhalation; therefore, the inspired air mixes with this residual air and lowers the partial pressure of oxygen within the alveoli. This means that there is a lower concentration of oxygen in the lungs than is found in the air outside the body. Knowing the RQ, the partial pressure of oxygen in the alveoli can be calculated:

$$\text{alveolar } P_{O_2} = \text{inspired } P_{O_2} - \left(\frac{\text{alveolar } P_{O_2}}{RQ}\right)$$

With an RQ of 0.8 and a P_{CO_2} in the alveoli of 40 mm Hg, the alveolar P_{O_2} is equal to:

$$\text{alveolar } P_{O_2} = 150 \text{ mm Hg} - \left(\frac{40 \text{ mm Hg}}{0.8}\right) = 100 \text{ mm Hg}.$$

Notice that this pressure is less than the external air. Therefore, the oxygen will flow from the inspired air in the lung ($P_{O_2} = 150$ mm Hg) into the bloodstream ($P_{O_2} = 100$ mm Hg) ([link]).

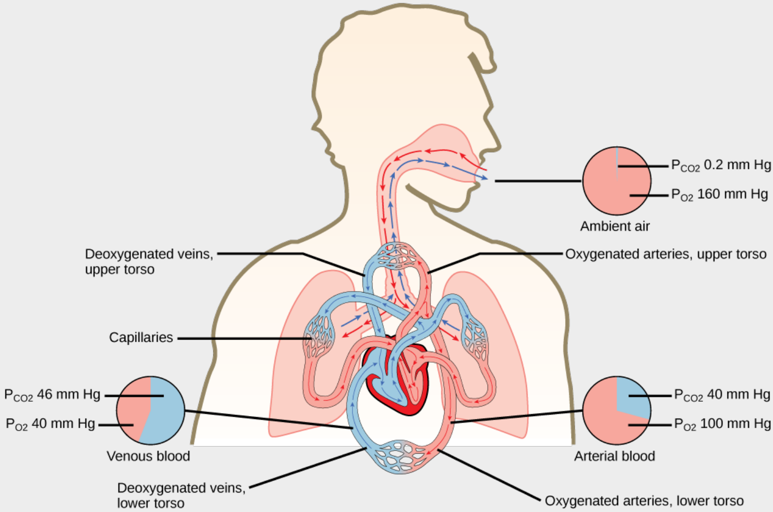
In the lungs, oxygen diffuses out of the alveoli and into the capillaries surrounding the alveoli. Oxygen (about 98 percent) binds reversibly to the respiratory pigment hemoglobin found in red blood cells (RBCs). RBCs carry oxygen to the tissues where oxygen dissociates from the hemoglobin and diffuses into the cells of the tissues. More specifically, alveolar P_{O_2} is higher in the alveoli ($P_{ALVO_2} = 100$ mm Hg) than blood P_{O_2} (40 mm Hg) in the capillaries. Because this pressure gradient exists, oxygen diffuses down its pressure gradient, moving out of the alveoli and entering the blood of the capillaries where O_2 binds to hemoglobin. At the same time, alveolar P_{CO_2} is lower $P_{ALVO_2} = 40$ mm Hg than blood $P_{CO_2} =$ (45 mm Hg). CO_2 diffuses down its pressure gradient, moving out of the capillaries and entering the alveoli.

Oxygen and carbon dioxide move independently of each other; they diffuse down their own pressure gradients. As blood leaves the lungs through the

pulmonary veins, the **venous P_{O_2}** $= 100$ mm Hg, whereas the **venous P_{CO_2}** $= 40$ mm Hg. As blood enters the systemic capillaries, the blood will lose oxygen and gain carbon dioxide because of the pressure difference of the tissues and blood. In systemic capillaries, $P_{O_2} = 100$ mm Hg, but in the tissue cells, $P_{O_2} = 40$ mm Hg. This pressure gradient drives the diffusion of oxygen out of the capillaries and into the tissue cells. At the same time, blood $P_{CO_2} = 40$ mm Hg and systemic tissue $P_{CO_2} = 45$ mm Hg. The pressure gradient drives CO_2 out of tissue cells and into the capillaries. The blood returning to the lungs through the pulmonary arteries has a venous $P_{O_2} = 40$ mm Hg and a $P_{CO_2} = 45$ mm Hg. The blood enters the lung capillaries where the process of exchanging gases between the capillaries and alveoli begins again ([link]).

Visual Connection

The partial pressures of oxygen and carbon dioxide change as blood moves through the body.



Which of the following statements is false?

1. In the tissues, P_{O_2} drops as blood passes from the arteries to the veins, while P_{CO_2} increases.
2. Blood travels from the lungs to the heart to body tissues, then back to the heart, then the lungs.
3. Blood travels from the lungs to the heart to body tissues, then back to the lungs, then the heart.
4. P_{O_2} is higher in air than in the lungs.

In short, the change in partial pressure from the alveoli to the capillaries drives the oxygen into the tissues and the carbon dioxide into the blood from the tissues. The blood is then transported to the lungs where differences in pressure in the alveoli

result in the movement of carbon dioxide out of the blood into the lungs, and oxygen into the blood.

Link to Learning

Watch this video to learn how to carry out spirometry.

https://www.openstax.org/l/spirometry

Section Summary

The lungs can hold a large volume of air, but they are not usually filled to maximal capacity. Lung volume measurements include tidal volume, expiratory reserve volume, inspiratory reserve volume, and residual volume. The sum of these equals the total lung capacity. Gas movement into or out of the lungs is dependent on the pressure of the gas. Air is a mixture of gases; therefore, the partial pressure of each gas can be calculated to determine how the gas will flow in the lung. The difference between the partial pressure of the gas in the air drives oxygen into the tissues and carbon dioxide out of the body.

Visual Connection Questions

[link] Which of the following statements is false?

1. In the tissues, P_{O_2} drops as blood passes from the arteries to the veins, while P_{CO_2} increases.
2. Blood travels from the lungs to the heart to body tissues, then back to the heart, then the lungs.
3. Blood travels from the lungs to the heart to body tissues, then back to the lungs, then the heart.
4. P_{O_2} is higher in air than in the lungs.

[link] C

Review Questions

The inspiratory reserve volume measures the _____.

1. amount of air remaining in the lung after a maximal exhalation
2. amount of air that the lung holds

3. amount of air that can be further exhaled after a normal breath
4. amount of air that can be further inhaled after a normal breath

---

D

Of the following, which does not explain why the partial pressure of oxygen is lower in the lung than in the external air?

1. Air in the lung is humidified; therefore, water vapor pressure alters the pressure.
2. Carbon dioxide mixes with oxygen.
3. Oxygen is moved into the blood and is headed to the tissues.
4. Lungs exert a pressure on the air to reduce the oxygen pressure.

D

The total lung capacity is calculated using which of the following formulas?

1. residual volume + tidal volume + inspiratory reserve volume
2. residual volume + expiratory reserve volume + inspiratory reserve volume

3. expiratory reserve volume + tidal volume + inspiratory reserve volume
4. residual volume + expiratory reserve volume + tidal volume + inspiratory reserve volume

---

D

Critical Thinking Questions

What does FEV1/FVC measure? What factors may affect FEV1/FVC?

FEV1/FVC measures the forced expiratory volume in one second in relation to the total forced vital capacity (the total amount of air that is exhaled from the lung from a maximal inhalation). This ratio changes with alterations in lung function that arise from diseases such as fibrosis, asthma, and COPD.

What is the reason for having residual volume in the lung?

If all the air in the lung were exhaled, then

opening the alveoli for the next inspiration would be very difficult. This is because the tissues would stick together.

How can a decrease in the percent of oxygen in the air affect the movement of oxygen in the body?

Oxygen moves from the lung to the bloodstream to the tissues according to the pressure gradient. This is measured as the partial pressure of oxygen. If the amount of oxygen drops in the inspired air, there would be reduced partial pressure. This would decrease the driving force that moves the oxygen into the blood and into the tissues. P_{O_2} is also reduced at high elevations: P_{O_2} at high elevations is lower than at sea level because the total atmospheric pressure is less than atmospheric pressure at sea level.

If a patient has increased resistance in his or her lungs, how can this be detected by a doctor? What does this mean?

A doctor can detect a restrictive disease using spirometry. By detecting the rate at which air can be expelled from the lung, a diagnosis of

fibrosis or another restrictive disease can be made.

Glossary

alveolar P_{O_2}
: partial pressure of oxygen in the alveoli (usually around 100 mmHg)

expiratory reserve volume (ERV)
: amount of additional air that can be exhaled after a normal exhalation

FEV1/FVC ratio
: ratio of how much air can be forced out of the lung in one second to the total amount that is forced out of the lung; a measurement of lung function that can be used to detect disease states

forced expiratory volume (FEV)
: (also, forced vital capacity) measure of how much air can be forced out of the lung from maximal inspiration over a specific amount of time

functional residual capacity (FRC)
: expiratory reserve volume plus residual volume

inspiratory capacity (IC)

tidal volume plus inspiratory reserve volume

inspiratory reserve volume (IRV)
amount of additional air that can be inspired after a normal inhalation

lung capacity
measurement of two or more lung volumes (how much air can be inhaled from the end of an expiration to maximal capacity)

lung volume
measurement of air for one lung function (normal inhalation or exhalation)

partial pressure
amount of pressure exerted by one gas within a mixture of gases

residual volume (RV)
amount of air remaining in the lung after a maximal expiration

respiratory quotient (RQ)
ratio of carbon dioxide production to each oxygen molecule consumed

spirometry
method to measure lung volumes and to diagnose lung diseases

tidal volume (TV)
amount of air that is inspired and expired

during normal breathing

total lung capacity (TLC)
: sum of the residual volume, expiratory reserve volume, tidal volume, and inspiratory reserve volume

venous P CO 2
: partial pressure of carbon dioxide in the veins (40 mm Hg in the pulmonary veins)

venous P O 2
: partial pressure of oxygen in the veins (100 mm Hg in the pulmonary veins)

vital capacity (VC)
: sum of the expiratory reserve volume, tidal volume, and inspiratory reserve volume

Breathing
By the end of this section, you will be able to do the following:

- Describe how the structures of the lungs and thoracic cavity control the mechanics of breathing
- Explain the importance of compliance and resistance in the lungs
- Discuss problems that may arise due to a V/Q mismatch

Mammalian lungs are located in the thoracic cavity where they are surrounded and protected by the rib cage, intercostal muscles, and bound by the chest wall. The bottom of the lungs is contained by the diaphragm, a skeletal muscle that facilitates breathing. Breathing requires the coordination of the lungs, the chest wall, and most importantly, the diaphragm.

Types of Breathing

Amphibians have evolved multiple ways of breathing. Young amphibians, like tadpoles, use gills to breathe, and they don't leave the water. Some amphibians retain gills for life. As the tadpole grows, the gills disappear and lungs grow. These lungs are primitive and not as evolved as

mammalian lungs. Adult amphibians are lacking or have a reduced diaphragm, so breathing via lungs is forced. The other means of breathing for amphibians is diffusion across the skin. To aid this diffusion, amphibian skin must remain moist.

Birds face a unique challenge with respect to breathing: They fly. Flying consumes a great amount of energy; therefore, birds require a lot of oxygen to aid their metabolic processes. Birds have evolved a respiratory system that supplies them with the oxygen needed to enable flying. Similar to mammals, birds have lungs, which are organs specialized for gas exchange. Oxygenated air, taken in during inhalation, diffuses across the surface of the lungs into the bloodstream, and carbon dioxide diffuses from the blood into the lungs and expelled during exhalation. The details of breathing between birds and mammals differ substantially.

In addition to lungs, birds have air sacs inside their body. Air flows in one direction from the posterior air sacs to the lungs and out of the anterior air sacs. The flow of air is in the opposite direction from blood flow, and gas exchange takes place much more efficiently. This type of breathing enables birds to obtain the requisite oxygen, even at higher altitudes where the oxygen concentration is low. This directionality of airflow requires two cycles of air intake and exhalation to completely get the air out of the lungs.

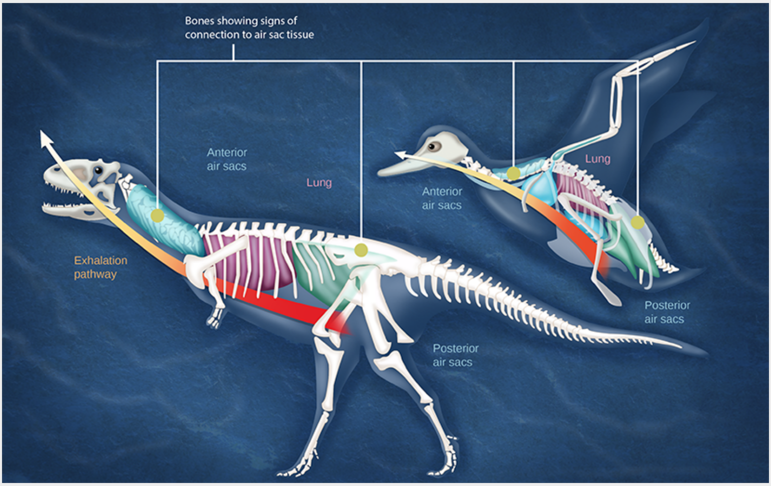
Evolution Connection

Avian Respiration

Birds have evolved a respiratory system that enables them to fly. Flying is a high-energy process and requires a lot of oxygen. Furthermore, many birds fly in high altitudes where the concentration of oxygen is low. How did birds evolve a respiratory system that is so unique?

Decades of research by paleontologists have shown that birds evolved from therapods, meat-eating dinosaurs ([link]). In fact, fossil evidence shows that meat-eating dinosaurs that lived more than 100 million years ago had a similar flow-through respiratory system with lungs and air sacs. *Archaeopteryx* and *Xiaotingia*, for example, were flying dinosaurs and are believed to be early precursors of birds.

Dinosaurs, from which birds descended, have similar hollow bones and are believed to have had a similar respiratory system. (credit b: modification of work by Zina Deretsky, National Science Foundation)



Most of us consider that dinosaurs are extinct. However, modern birds are descendants of avian dinosaurs. The respiratory system of modern birds has been evolving for hundreds of millions of years.

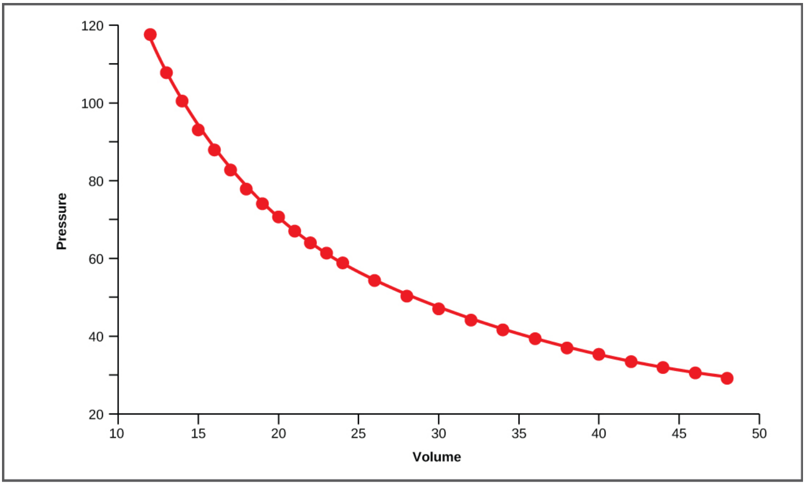
All mammals have lungs that are the main organs for breathing. Lung capacity has evolved to support the animal's activities. During inhalation, the lungs expand with air, and oxygen diffuses across the lung's surface and enters the bloodstream. During exhalation, the lungs expel air and lung volume decreases. In the next few sections, the process of human breathing will be explained.

This graph shows data from Boyle's original 1662 experiment, which shows that pressure and volume are inversely related. No units are given as Boyle used arbitrary units in his experiments. The lungs, chest wall, and diaphragm are all involved in

respiration, both (a) inhalation and (b) expiration. (credit: modification of work by Mariana Ruiz Villareal) A tissue layer called pleura surrounds the lung and interior of the thoracic cavity. (credit: modification of work by NCI)

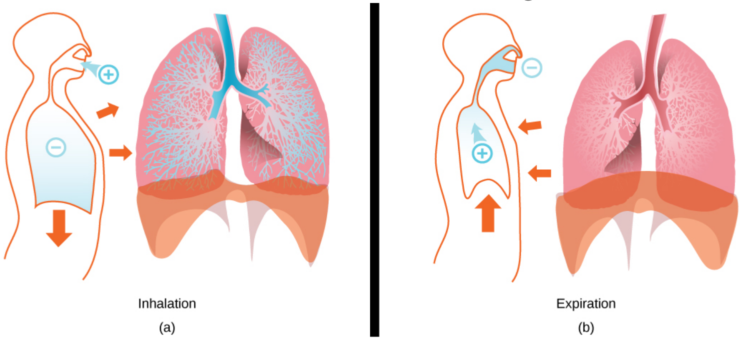
The Mechanics of Human Breathing

Boyle's Law is the gas law that states that in a closed space, pressure and volume are inversely related. As volume decreases, pressure increases and vice versa ([link]). The relationship between gas pressure and volume helps to explain the mechanics of breathing.



There is always a slightly negative pressure within the thoracic cavity, which aids in keeping the airways of the lungs open. During inhalation, volume increases as a result of contraction of the diaphragm, and pressure decreases (according to Boyle's Law). This decrease of pressure in the

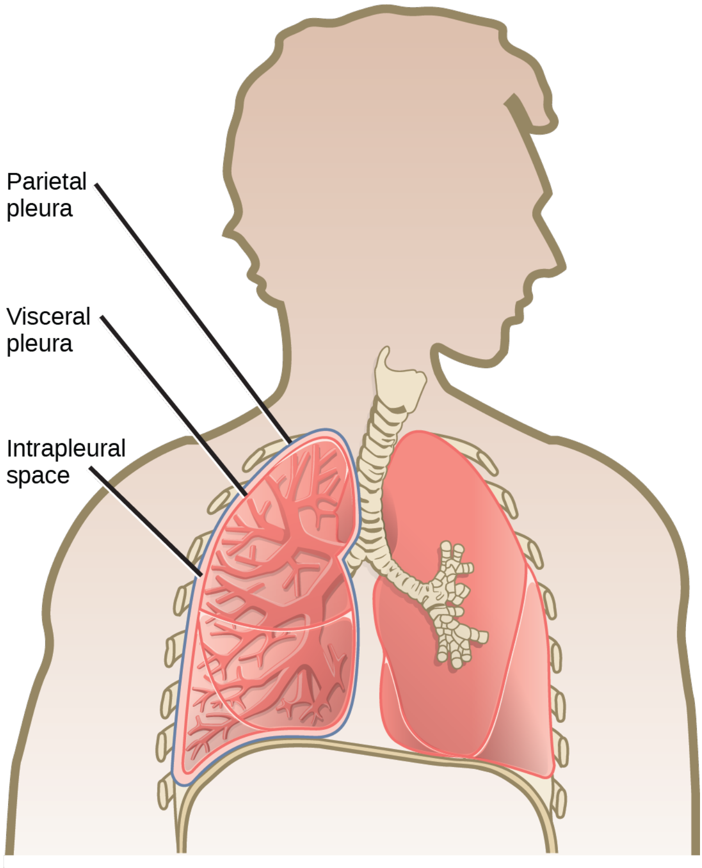
thoracic cavity relative to the environment makes the cavity less than the atmosphere ([link]**a**). Because of this drop in pressure, air rushes into the respiratory passages. To increase the volume of the lungs, the chest wall expands. This results from the contraction of the **intercostal muscles**, the muscles that are connected to the rib cage. Lung volume expands because the diaphragm contracts and the intercostal muscles contract, thus expanding the thoracic cavity. This increase in the volume of the thoracic cavity lowers pressure compared to the atmosphere, so air rushes into the lungs, thus increasing its volume. The resulting increase in volume is largely attributed to an increase in alveolar space, because the bronchioles and bronchi are stiff structures that do not change in size.



The chest wall expands out and away from the lungs. The lungs are elastic; therefore, when air fills the lungs, the **elastic recoil** within the tissues of the lung exerts pressure back toward the interior of the lungs. These outward and inward forces compete to inflate and deflate the lung with every breath. Upon exhalation, the lungs recoil to force the air out of

the lungs, and the intercostal muscles relax, returning the chest wall back to its original position ([link]**b**). The diaphragm also relaxes and moves higher into the thoracic cavity. This increases the pressure within the thoracic cavity relative to the environment, and air rushes out of the lungs. The movement of air out of the lungs is a passive event. No muscles are contracting to expel the air.

Each lung is surrounded by an invaginated sac. The layer of tissue that covers the lung and dips into spaces is called the visceral **pleura**. A second layer of parietal pleura lines the interior of the thorax ([link]). The space between these layers, the **intrapleural space**, contains a small amount of fluid that protects the tissue and reduces the friction generated from rubbing the tissue layers together as the lungs contract and relax. **Pleurisy** results when these layers of tissue become inflamed; it is painful because the inflammation increases the pressure within the thoracic cavity and reduces the volume of the lung.



Link to Learning

View how Boyle's Law is related to breathing and watch a video on Boyle's Law.
https://www.openstax.org/l/boyle_breathing

The ratio of FEV1 (the amount of air that can be forcibly exhaled in one second after taking a deep

breath) to FVC (the total amount of air that can be forcibly exhaled) can be used to diagnose whether a person has restrictive or obstructive lung disease. In restrictive lung disease, FVC is reduced but airways are not obstructed, so the person is able to expel air reasonably fast. In obstructive lung disease, airway obstruction results in slow exhalation as well as reduced FVC. Thus, the FEV1/FVC ratio is lower in persons with obstructive lung disease (less than 69 percent) than in persons with restrictive disease (88 to 90 percent).

The Work of Breathing

The number of breaths per minute is the **respiratory rate**. On average, under non-exertion conditions, the human respiratory rate is 12–15 breaths/minute. The respiratory rate contributes to the **alveolar ventilation**, or how much air moves into and out of the alveoli. Alveolar ventilation prevents carbon dioxide buildup in the alveoli. There are two ways to keep the alveolar ventilation constant: increase the respiratory rate while decreasing the tidal volume of air per breath (shallow breathing), or decrease the respiratory rate while increasing the tidal volume per breath. In either case, the ventilation remains the same, but the work done and type of work needed are quite different. Both tidal volume and respiratory rate are closely regulated when oxygen demand increases.

There are two types of work conducted during respiration, flow-resistive and elastic work. **Flow-resistive** refers to the work of the alveoli and tissues in the lung, whereas **elastic work** refers to the work of the intercostal muscles, chest wall, and diaphragm. Increasing the respiration rate increases the flow-resistive work of the airways and decreases the elastic work of the muscles. Decreasing the respiratory rate reverses the type of work required.

Surfactant

The air-tissue/water interface of the alveoli has a high surface tension. This surface tension is similar to the surface tension of water at the liquid-air interface of a water droplet that results in the bonding of the water molecules together. **Surfactant** is a complex mixture of phospholipids and lipoproteins that works to reduce the surface tension that exists between the alveoli tissue and the air found within the alveoli. By lowering the surface tension of the alveolar fluid, it reduces the tendency of alveoli to collapse.

Surfactant works like a detergent to reduce the surface tension and allows for easier inflation of the airways. When a balloon is first inflated, it takes a large amount of effort to stretch the plastic and start to inflate the balloon. If a little bit of detergent was applied to the interior of the balloon, then the amount of effort or work needed to begin to inflate

the balloon would decrease, and it would become much easier to start blowing up the balloon. This same principle applies to the airways. A small amount of surfactant to the airway tissues reduces the effort or work needed to inflate those airways. Babies born prematurely sometimes do not produce enough surfactant. As a result, they suffer from **respiratory distress syndrome**, because it requires more effort to inflate their lungs. Surfactant is also important for preventing collapse of small alveoli relative to large alveoli.

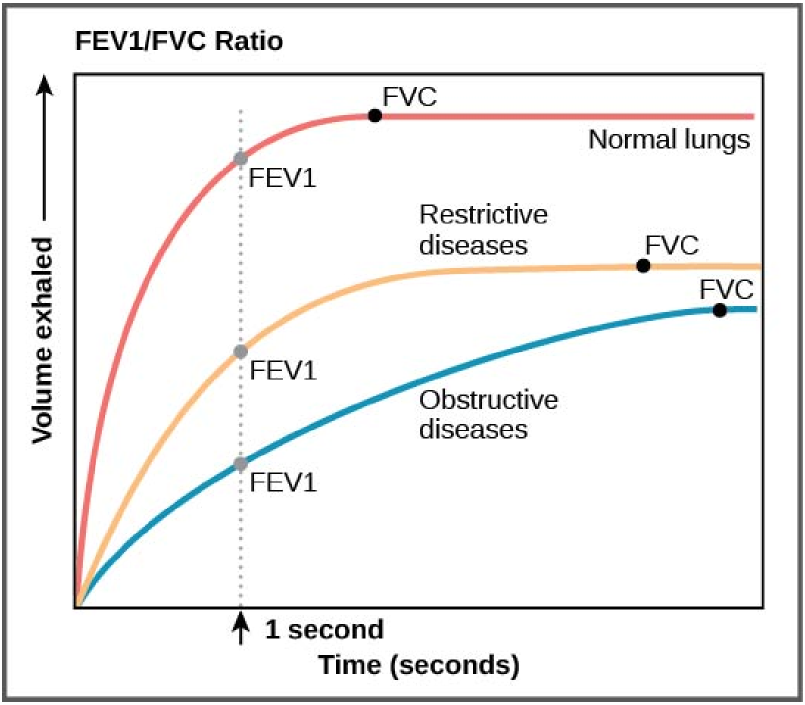
Lung Resistance and Compliance

Pulmonary diseases reduce the rate of gas exchange into and out of the lungs. Two main causes of decreased gas exchange are **compliance** (how elastic the lung is) and **resistance** (how much obstruction exists in the airways). A change in either can dramatically alter breathing and the ability to take in oxygen and release carbon dioxide.

Examples of **restrictive diseases** are respiratory distress syndrome and pulmonary fibrosis. In both diseases, the airways are less compliant and they are stiff or fibrotic. There is a decrease in compliance because the lung tissue cannot bend and move. In these types of restrictive diseases, the intrapleural pressure is more positive and the airways collapse upon exhalation, which traps air in the lungs. Forced or **functional vital capacity (FVC)**, which is

the amount of air that can be forcibly exhaled after taking the deepest breath possible, is much lower than in normal patients, and the time it takes to exhale most of the air is greatly prolonged ([link]). A patient suffering from these diseases cannot exhale the normal amount of air.

Obstructive diseases and conditions include emphysema, asthma, and pulmonary edema. In emphysema, which mostly arises from smoking tobacco, the walls of the alveoli are destroyed, decreasing the surface area for gas exchange. The overall compliance of the lungs is increased, because as the alveolar walls are damaged, lung elastic recoil decreases due to a loss of elastic fibers, and more air is trapped in the lungs at the end of exhalation. Asthma is a disease in which inflammation is triggered by environmental factors. Inflammation obstructs the airways. The obstruction may be due to edema (fluid accumulation), smooth muscle spasms in the walls of the bronchioles, increased mucus secretion, damage to the epithelia of the airways, or a combination of these events. Those with asthma or edema experience increased occlusion from increased inflammation of the airways. This tends to block the airways, preventing the proper movement of gases ([link]). Those with obstructive diseases have large volumes of air trapped after exhalation and breathe at a very high lung volume to compensate for the lack of airway recruitment.



Dead Space: V/Q Mismatch

Pulmonary circulation pressure is very low compared to that of the systemic circulation. It is also independent of cardiac output. This is because of a phenomenon called **recruitment**, which is the process of opening airways that normally remain closed when cardiac output increases. As cardiac output increases, the number of capillaries and arteries that are perfused (filled with blood) increases. These capillaries and arteries are not always in use but are ready if needed. At times, however, there is a mismatch between the amount

of air (ventilation, V) and the amount of blood (perfusion, Q) in the lungs. This is referred to as **ventilation/perfusion (V/Q) mismatch**.

There are two types of V/Q mismatch. Both produce **dead space**, regions of broken down or blocked lung tissue. Dead spaces can severely impact breathing, because they reduce the surface area available for gas diffusion. As a result, the amount of oxygen in the blood decreases, whereas the carbon dioxide level increases. Dead space is created when no ventilation and/or perfusion takes place. **Anatomical dead space** or anatomical shunt, arises from an anatomical failure, while **physiological dead space** or physiological shunt, arises from a functional impairment of the lung or arteries.

An example of an anatomical shunt is the effect of gravity on the lungs. The lung is particularly susceptible to changes in the magnitude and direction of gravitational forces. When someone is standing or sitting upright, the pleural pressure gradient leads to increased ventilation further down in the lung. As a result, the intrapleural pressure is more negative at the base of the lung than at the top, and more air fills the bottom of the lung than the top. Likewise, it takes less energy to pump blood to the bottom of the lung than to the top when in a prone position. Perfusion of the lung is not uniform while standing or sitting. This is a result of hydrostatic forces combined with the effect of

airway pressure. An anatomical shunt develops because the ventilation of the airways does not match the perfusion of the arteries surrounding those airways. As a result, the rate of gas exchange is reduced. Note that this does not occur when lying down, because in this position, gravity does not preferentially pull the bottom of the lung down.

A physiological shunt can develop if there is infection or edema in the lung that obstructs an area. This will decrease ventilation but not affect perfusion; therefore, the V/Q ratio changes and gas exchange is affected.

The lung can compensate for these mismatches in ventilation and perfusion. If ventilation is greater than perfusion, the arterioles dilate and the bronchioles constrict. This increases perfusion and reduces ventilation. Likewise, if ventilation is less than perfusion, the arterioles constrict and the bronchioles dilate to correct the imbalance.

Link to Learning

View the mechanics of breathing.

https://www.openstax.org/l/breathing

Section Summary

The structure of the lungs and thoracic cavity control the mechanics of breathing. Upon inspiration, the diaphragm contracts and lowers. The intercostal muscles contract and expand the chest wall outward. The intrapleural pressure drops, the lungs expand, and air is drawn into the airways. When exhaling, the intercostal muscles and diaphragm relax, returning the intrapleural pressure back to the resting state. The lungs recoil and airways close. The air passively exits the lung. There is high surface tension at the air-airway interface in the lung. Surfactant, a mixture of phospholipids and lipoproteins, acts like a detergent in the airways to reduce surface tension and allow for opening of the alveoli.

Breathing and gas exchange are both altered by changes in the compliance and resistance of the lung. If the compliance of the lung decreases, as occurs in restrictive diseases like fibrosis, the airways stiffen and collapse upon exhalation. Air becomes trapped in the lungs, making breathing more difficult. If resistance increases, as happens with asthma or emphysema, the airways become obstructed, trapping air in the lungs and causing breathing to become difficult. Alterations in the ventilation of the airways or perfusion of the arteries can affect gas exchange. These changes in ventilation and perfusion, called V/Q mismatch, can

arise from anatomical or physiological changes.

Review Questions

How would paralysis of the diaphragm alter inspiration?

1. It would prevent contraction of the intercostal muscles.
2. It would prevent inhalation because the intrapleural pressure would not change.
3. It would decrease the intrapleural pressure and allow more air to enter the lungs.
4. It would slow expiration because the lung would not relax.

B

Restrictive airway diseases ______.

1. increase the compliance of the lung
2. decrease the compliance of the lung
3. increase the lung volume
4. decrease the work of breathing

B

Alveolar ventilation remains constant when _____.

1. the respiratory rate is increased while the volume of air per breath is decreased
2. the respiratory rate and the volume of air per breath are increased
3. the respiratory rate is decreased while increasing the volume per breath
4. both a and c

D

Critical Thinking Questions

How would increased airway resistance affect intrapleural pressure during inhalation?

Increased airway resistance increases the volume and pressure in the lung; therefore, the intrapleural pressure would be less negative and breathing would be more difficult.

Explain how a puncture to the thoracic cavity

(from a knife wound, for instance) could alter the ability to inhale.

A puncture to the thoracic cavity would equalize the pressure inside the thoracic cavity to the outside environment. For the lung to function properly, the intrapleural pressure must be negative. This is caused by the contraction of the diaphragm pulling the lungs down and drawing air into the lungs.

When someone is standing, gravity stretches the bottom of the lung down toward the floor to a greater extent than the top of the lung. What implication could this have on the flow of air in the lungs? Where does gas exchange occur in the lungs?

The lung is particularly susceptible to changes in the magnitude and direction of gravitational forces. When someone is standing or sitting upright, the pleural pressure gradient leads to increased ventilation further down in the lung.

Glossary

alveolar ventilation
: how much air is in the alveoli

anatomical dead space
: (also, anatomical shunt) region of the lung that lacks proper ventilation/perfusion due to an anatomical block

compliance
: measurement of the elasticity of the lung

dead space
: area in the lung that lacks proper ventilation or perfusion

elastic recoil
: property of the lung that drives the lung tissue inward

elastic work
: work conducted by the intercostal muscles, chest wall, and diaphragm

flow-resistive
: work of breathing performed by the alveoli and tissues in the lung

functional vital capacity (FVC)
: amount of air that can be forcibly exhaled after taking the deepest breath possible

intercostal muscle
: muscle connected to the rib cage that contracts upon inspiration

intrapleural space

space between the layers of pleura

obstructive disease
: disease (such as emphysema and asthma) that arises from obstruction of the airways; compliance increases in these diseases

physiological dead space
: (also, physiological shunt) region of the lung that lacks proper ventilation/perfusion due to a physiological change in the lung (like inflammation or edema)

pleura
: tissue layer that surrounds the lungs and lines the interior of the thoracic cavity

pleurisy
: painful inflammation of the pleural tissue layers

recruitment
: process of opening airways that normally remain closed when the cardiac output increases

resistance
: measurement of lung obstruction

respiratory distress syndrome
: disease that arises from a deficient amount of surfactant

respiratory rate
: number of breaths per minute

restrictive disease
: disease that results from a restriction and decreased compliance of the alveoli; respiratory distress syndrome and pulmonary fibrosis are examples

surfactant
: detergent-like liquid in the airways that lowers the surface tension of the alveoli to allow for expansion

ventilation/perfusion (V/Q) mismatch
: region of the lung that lacks proper alveolar ventilation (V) and/or arterial perfusion (Q)

Transport of Gases in Human Bodily Fluids
By the end of this section, you will be able to do the following:

- Describe how oxygen is bound to hemoglobin and transported to body tissues
- Explain how carbon dioxide is transported from body tissues to the lungs

Once the oxygen diffuses across the alveoli, it enters the bloodstream and is transported to the tissues where it is unloaded, and carbon dioxide diffuses out of the blood and into the alveoli to be expelled from the body. Although gas exchange is a continuous process, the oxygen and carbon dioxide are transported by different mechanisms.
The protein inside (a) red blood cells that carries oxygen to cells and carbon dioxide to the lungs is (b) hemoglobin. Hemoglobin is made up of four symmetrical subunits and four heme groups. Iron associated with the heme binds oxygen. It is the iron in hemoglobin that gives blood its red color. Individuals with sickle cell anemia have crescent-shaped red blood cells. (credit: modification of work by Ed Uthman; scale-bar data from Matt Russell)

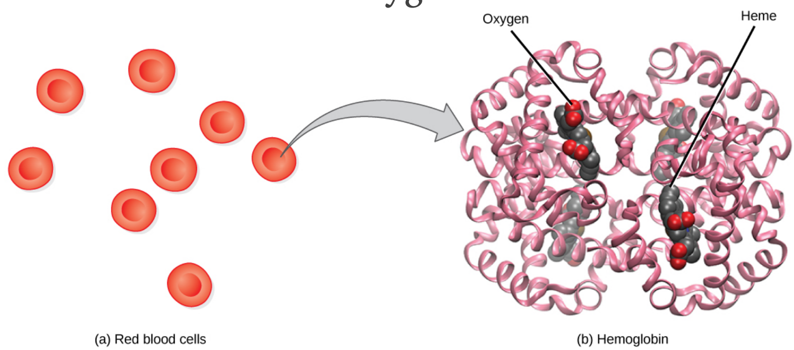
Transport of Oxygen in the Blood

Although oxygen dissolves in blood, only a small amount of oxygen is transported this way. Only 1.5

percent of oxygen in the blood is dissolved directly into the blood itself. Most oxygen—98.5 percent—is bound to a protein called hemoglobin and carried to the tissues.

Hemoglobin

Hemoglobin, or Hb, is a protein molecule found in red blood cells (erythrocytes) made of four subunits: two alpha subunits and two beta subunits ([link]). Each subunit surrounds a central **heme group** that contains iron and binds one oxygen molecule, allowing each hemoglobin molecule to bind four oxygen molecules. Molecules with more oxygen bound to the heme groups are brighter red. As a result, oxygenated arterial blood where the Hb is carrying four oxygen molecules is bright red, while venous blood that is deoxygenated is darker red.

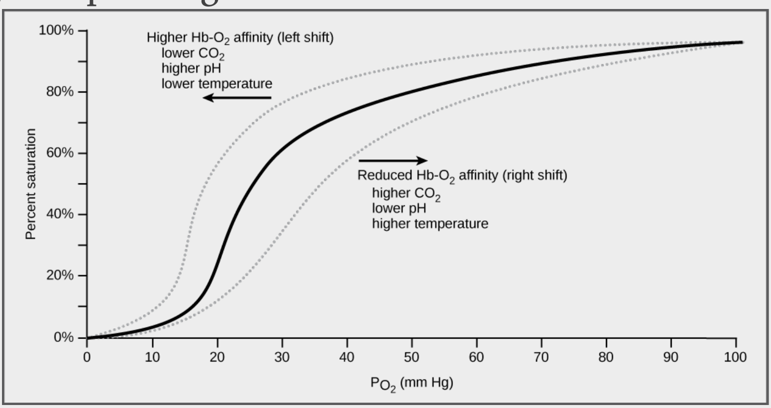


It is easier to bind a second and third oxygen molecule to Hb than the first molecule. This is because the hemoglobin molecule changes its shape, or conformation, as oxygen binds. The fourth

oxygen is then more difficult to bind. The binding of oxygen to hemoglobin can be plotted as a function of the partial pressure of oxygen in the blood (x-axis) versus the relative Hb-oxygen saturation (y-axis). The resulting graph—an **oxygen dissociation curve**—is sigmoidal, or S-shaped ([link]). As the partial pressure of oxygen increases, the hemoglobin becomes increasingly saturated with oxygen.

Visual Connection

The oxygen dissociation curve demonstrates that, as the partial pressure of oxygen increases, more oxygen binds hemoglobin. However, the affinity of hemoglobin for oxygen may shift to the left or the right depending on environmental conditions.



The kidneys are responsible for removing excess H + ions from the blood. If the kidneys fail, what would happen to blood pH and to hemoglobin affinity for oxygen?

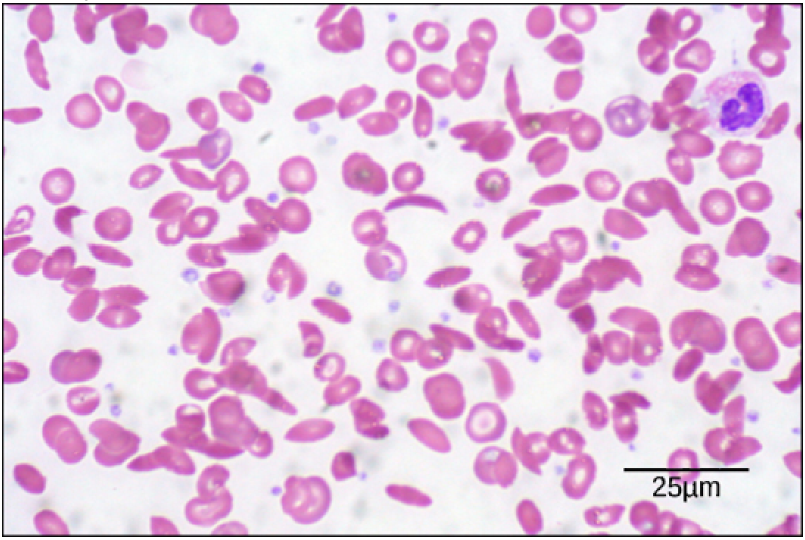
Factors That Affect Oxygen Binding

The **oxygen-carrying capacity** of hemoglobin determines how much oxygen is carried in the blood. In addition to P_{O_2}, other environmental factors and diseases can affect oxygen carrying capacity and delivery.

Carbon dioxide levels, blood pH, and body temperature affect oxygen-carrying capacity ([link]). When carbon dioxide is in the blood, it reacts with water to form bicarbonate (HCO_3^-) and hydrogen ions (H^+). As the level of carbon dioxide in the blood increases, more H^+ is produced and the pH decreases. This increase in carbon dioxide and subsequent decrease in pH reduce the affinity of hemoglobin for oxygen. The oxygen dissociates from the Hb molecule, shifting the oxygen dissociation curve to the right. Therefore, more oxygen is needed to reach the same hemoglobin saturation level as when the pH was higher. A similar shift in the curve also results from an increase in body temperature. Increased temperature, such as from increased activity of skeletal muscle, causes the affinity of hemoglobin for oxygen to be reduced.

Diseases like sickle cell anemia and thalassemia decrease the blood's ability to deliver oxygen to tissues and its oxygen-carrying capacity. In **sickle cell anemia**, the shape of the red blood cell is

crescent-shaped, elongated, and stiffened, reducing its ability to deliver oxygen ([link]). In this form, red blood cells cannot pass through the capillaries. This is painful when it occurs. **Thalassemia** is a rare genetic disease caused by a defect in either the alpha or the beta subunit of Hb. Patients with thalassemia produce a high number of red blood cells, but these cells have lower-than-normal levels of hemoglobin. Therefore, the oxygen-carrying capacity is diminished.



As percent CO increases, the oxygen saturation of hemoglobin decreases.

Transport of Carbon Dioxide in the Blood

Carbon dioxide molecules are transported in the blood from body tissues to the lungs by one of three methods: dissolution directly into the blood, binding to hemoglobin, or carried as a bicarbonate ion.

Several properties of carbon dioxide in the blood affect its transport. First, carbon dioxide is more soluble in blood than oxygen. About 5 to 7 percent of all carbon dioxide is dissolved in the plasma. Second, carbon dioxide can bind to plasma proteins or can enter red blood cells and bind to hemoglobin. This form transports about 10 percent of the carbon dioxide. When carbon dioxide binds to hemoglobin, a molecule called **carbaminohemoglobin** is formed. Binding of carbon dioxide to hemoglobin is reversible. Therefore, when it reaches the lungs, the carbon dioxide can freely dissociate from the hemoglobin and be expelled from the body.

Third, the majority of carbon dioxide molecules (85 percent) are carried as part of the **bicarbonate buffer system**. In this system, carbon dioxide diffuses into the red blood cells. **Carbonic anhydrase (CA)** within the red blood cells quickly converts the carbon dioxide into carbonic acid (H_2CO_3). Carbonic acid is an unstable intermediate molecule that immediately dissociates into **bicarbonate ions (HCO_3^-)** and hydrogen (H^+) ions. Since carbon dioxide is quickly converted into bicarbonate ions, this reaction allows for the continued uptake of carbon dioxide into the blood down its concentration gradient. It also results in the production of H^+ ions. If too much H^+ is produced, it can alter blood pH. However, hemoglobin binds to the free H^+ ions and thus limits shifts in pH. The newly synthesized

bicarbonate ion is transported out of the red blood cell into the liquid component of the blood in exchange for a chloride ion (Cl-); this is called the **chloride shift**. When the blood reaches the lungs, the bicarbonate ion is transported back into the red blood cell in exchange for the chloride ion. The H^+ ion dissociates from the hemoglobin and binds to the bicarbonate ion. This produces the carbonic acid intermediate, which is converted back into carbon dioxide through the enzymatic action of CA. The carbon dioxide produced is expelled through the lungs during exhalation.

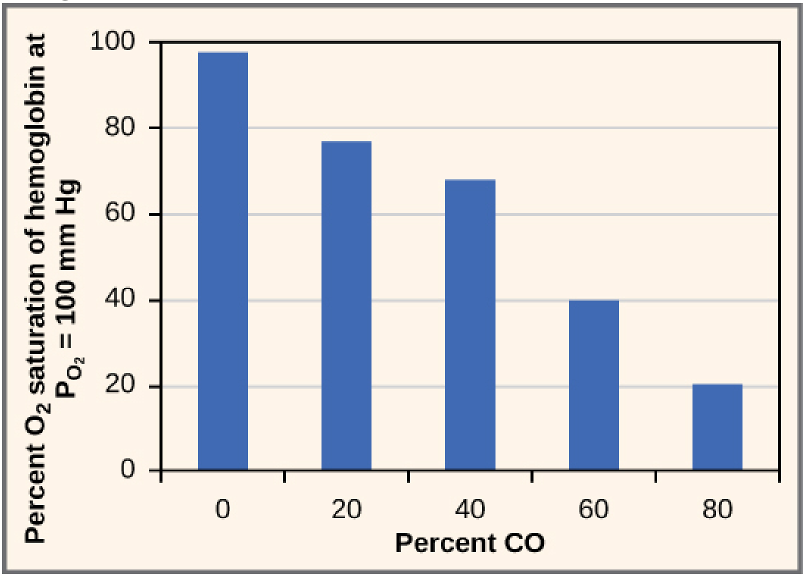
$CO_2 + H_2O \leftrightarrow H_2CO_3$ (carbonic acid) $\leftrightarrow$ $HCO_3^- + H^+$ (bicarbonate)

The benefit of the bicarbonate buffer system is that carbon dioxide is "soaked up" into the blood with little change to the pH of the system. This is important because it takes only a small change in the overall pH of the body for severe injury or death to result. The presence of this bicarbonate buffer system also allows for people to travel and live at high altitudes: When the partial pressure of oxygen and carbon dioxide change at high altitudes, the bicarbonate buffer system adjusts to regulate carbon dioxide while maintaining the correct pH in the body.

Carbon Monoxide Poisoning

While carbon dioxide can readily associate and

dissociate from hemoglobin, other molecules such as carbon monoxide (CO) cannot. Carbon monoxide has a greater affinity for hemoglobin than oxygen. Therefore, when carbon monoxide is present, it binds to hemoglobin preferentially over oxygen. As a result, oxygen cannot bind to hemoglobin, so very little oxygen is transported through the body ([link]). Carbon monoxide is a colorless, odorless gas and is therefore difficult to detect. It is produced by gas-powered vehicles and tools. Carbon monoxide can cause headaches, confusion, and nausea; long-term exposure can cause brain damage or death. Administering 100 percent (pure) oxygen is the usual treatment for carbon monoxide poisoning. Administration of pure oxygen speeds up the separation of carbon monoxide from hemoglobin.



Section Summary

Hemoglobin is a protein found in red blood cells that is comprised of two alpha and two beta subunits that surround an iron-containing heme group. Oxygen readily binds this heme group. The ability of oxygen to bind increases as more oxygen molecules are bound to heme. Disease states and altered conditions in the body can affect the binding ability of oxygen, and increase or decrease its ability to dissociate from hemoglobin.

Carbon dioxide can be transported through the blood via three methods. It is dissolved directly in the blood, bound to plasma proteins or hemoglobin, or converted into bicarbonate. The majority of carbon dioxide is transported as part of the bicarbonate system. Carbon dioxide diffuses into red blood cells. Inside, carbonic anhydrase converts carbon dioxide into carbonic acid (H_2CO_3), which is subsequently hydrolyzed into bicarbonate (HCO_3^-) and H^+. The H^+ ion binds to hemoglobin in red blood cells, and bicarbonate is transported out of the red blood cells in exchange for a chloride ion. This is called the chloride shift. Bicarbonate leaves the red blood cells and enters the blood plasma. In the lungs, bicarbonate is transported back into the red blood cells in exchange for chloride. The H^+ dissociates from hemoglobin and combines with bicarbonate to form carbonic acid with the help of carbonic anhydrase, which further catalyzes the

reaction to convert carbonic acid back into carbon dioxide and water. The carbon dioxide is then expelled from the lungs.

Visual Connection Questions

[link] The kidneys are responsible for removing excess $H+$ ions from the blood. If the kidneys fail, what would happen to blood pH and to hemoglobin affinity for oxygen?

[link] The blood pH will drop and hemoglobin affinity for oxygen will decrease.

Review Questions

Which of the following will NOT facilitate the transfer of oxygen to tissues?

1. decreased body temperature
2. decreased pH of the blood
3. increased carbon dioxide
4. increased exercise

A

The majority of carbon dioxide in the blood is transported by _____.

1. binding to hemoglobin
2. dissolution in the blood
3. conversion to bicarbonate
4. binding to plasma proteins

C

The majority of oxygen in the blood is transported by _____.

1. dissolution in the blood
2. being carried as bicarbonate ions
3. binding to blood plasma
4. binding to hemoglobin

D

Critical Thinking Questions

What would happen if no carbonic anhydrase

were present in red blood cells?

Without carbonic anhydrase, carbon dioxide would not be hydrolyzed into carbonic acid or bicarbonate. Therefore, very little carbon dioxide (only 15 percent) would be transported in the blood away from the tissues.

How does the administration of 100 percent oxygen save a patient from carbon monoxide poisoning? Why wouldn't giving carbon dioxide work?

Carbon monoxide has a higher affinity for hemoglobin than oxygen. This means that carbon monoxide will preferentially bind to hemoglobin over oxygen. Administration of 100 percent oxygen is an effective therapy because at that concentration, oxygen will displace the carbon monoxide from the hemoglobin.

Glossary

bicarbonate buffer system
: system in the blood that absorbs carbon dioxide and regulates pH levels

bicarbonate (HCO_3^-) ion

ion created when carbonic acid dissociates into H^+ and (HCO_3^-)

carbaminohemoglobin
: molecule that forms when carbon dioxide binds to hemoglobin

carbonic anhydrase (CA)
: enzyme that catalyzes carbon dioxide and water into carbonic acid

chloride shift
: exchange of chloride for bicarbonate into or out of the red blood cell

heme group
: centralized iron-containing group that is surrounded by the alpha and beta subunits of hemoglobin

hemoglobin
: molecule in red blood cells that can bind oxygen, carbon dioxide, and carbon monoxide

oxygen-carrying capacity
: amount of oxygen that can be transported in the blood

oxygen dissociation curve
: curve depicting the affinity of oxygen for hemoglobin

sickle cell anemia

genetic disorder that affects the shape of red blood cells, and their ability to transport oxygen and move through capillaries

thalassemia
: rare genetic disorder that results in mutation of the alpha or beta subunits of hemoglobin, creating smaller red blood cells with less hemoglobin

Introduction
class="introduction" Just as highway systems transport people and goods through a complex network, the circulatory system transports nutrients, gases, and wastes throughout the animal body. (credit: modification of work by Andrey Belenko)

Most animals are complex multicellular organisms that require a mechanism for transporting nutrients throughout their bodies and removing waste products. The circulatory system has evolved over time from simple diffusion through cells in the early evolution of animals to a complex network of blood vessels that reach all parts of the human body. This extensive network supplies the cells, tissues, and organs with oxygen and nutrients, and removes carbon dioxide and waste, which are byproducts of respiration.

At the core of the human circulatory system is the heart. The size of a clenched fist, the human heart is protected beneath the rib cage. Made of specialized

and unique cardiac muscle, it pumps blood throughout the body and to the heart itself. Heart contractions are driven by intrinsic electrical impulses that the brain and endocrine hormones help to regulate. Understanding the heart's basic anatomy and function is important to understanding the body's circulatory and respiratory systems.

Gas exchange is one essential function of the circulatory system. A circulatory system is not needed in organisms with no specialized respiratory organs because oxygen and carbon dioxide diffuse directly between their body tissues and the external environment. However, in organisms that possess lungs and gills, oxygen must be transported from these specialized respiratory organs to the body tissues via a circulatory system. Therefore, circulatory systems have had to evolve to accommodate the great diversity of body sizes and body types present among animals.

Overview of the Circulatory System
By the end of this section, you will be able to do the following:

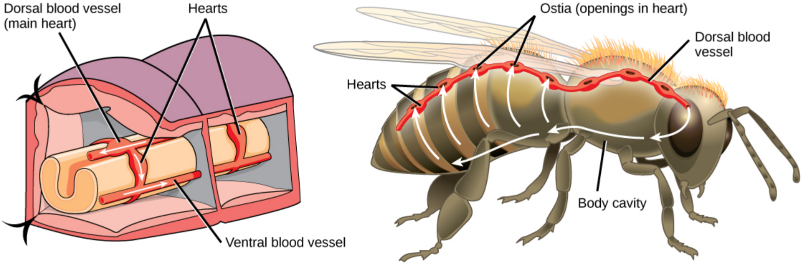
- Describe an open and closed circulatory system
- Describe interstitial fluid and hemolymph
- Compare and contrast the organization and evolution of the vertebrate circulatory system

In all animals, except a few simple types, the circulatory system is used to transport nutrients and gases through the body. Simple diffusion allows some water, nutrient, waste, and gas exchange into primitive animals that are only a few cell layers thick; however, bulk flow is the only method by which the entire body of larger more complex organisms is accessed.
In (a) closed circulatory systems, the heart pumps blood through vessels that are separate from the interstitial fluid of the body. Most vertebrates and some invertebrates, like this annelid earthworm, have a closed circulatory system. In (b) open circulatory systems, a fluid called hemolymph is pumped through a blood vessel that empties into the body cavity. Hemolymph returns to the blood vessel through openings called ostia. Arthropods like this bee and most mollusks have open circulatory systems.

Circulatory System Architecture

The circulatory system is effectively a network of cylindrical vessels: the arteries, veins, and capillaries that emanate from a pump, the heart. In all vertebrate organisms, as well as some invertebrates, this is a closed-loop system, in which the blood is not free in a cavity. In a **closed circulatory system**, blood is contained inside blood vessels and circulates **unidirectionally** from the heart around the systemic circulatory route, then returns to the heart again, as illustrated in [link]**a**. As opposed to a closed system, arthropods—including insects, crustaceans, and most mollusks—have an open circulatory system, as illustrated in [link]**b**. In an **open circulatory system**, the blood is not enclosed in the blood vessels but is pumped into a cavity called a **hemocoel** and is called **hemolymph** because the blood mixes with the **interstitial fluid**. As the heart beats and the animal moves, the hemolymph circulates around the organs within the body cavity and then reenters the hearts through openings called **ostia**. This movement allows for gas and nutrient exchange. An open circulatory system does not use as much energy as a closed system to operate or to maintain; however, there is a trade-off with the amount of blood that can be moved to metabolically active organs and tissues that require high levels of oxygen. In fact, one reason that insects with wing spans of up to two feet wide (70 cm) are not around today is probably because they were outcompeted by the arrival of birds 150 million years ago. Birds, having a closed

circulatory system, are thought to have moved more agilely, allowing them to get food faster and possibly to prey on the insects.



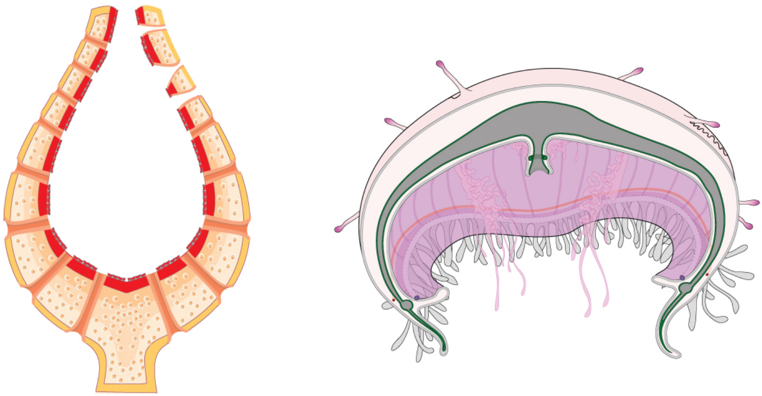
(a) Closed circulatory system (b) Open circulatory system

Simple animals consisting of a single cell layer such as the (a) sponge or only a few cell layers such as the (b) jellyfish do not have a circulatory system. Instead, gases, nutrients, and wastes are exchanged by diffusion. (a) Fish have the simplest circulatory systems of the vertebrates: blood flows unidirectionally from the two-chambered heart through the gills and then the rest of the body. (b) Amphibians have two circulatory routes: one for oxygenation of the blood through the lungs and skin, and the other to take oxygen to the rest of the body. The blood is pumped from a three-chambered heart with two atria and a single ventricle. (c) Reptiles also have two circulatory routes; however, blood is only oxygenated through the lungs. The heart is three chambered, but the ventricles are partially separated so some mixing of oxygenated and deoxygenated blood occurs except in crocodilians and birds. (d) Mammals and birds have the most efficient heart with four chambers that completely separate the oxygenated and deoxygenated blood; it pumps only oxygenated

blood through the body and deoxygenated blood to the lungs.

Circulatory System Variation in Animals

The circulatory system varies from simple systems in invertebrates to more complex systems in vertebrates. The simplest animals, such as the sponges (Porifera) and rotifers (Rotifera), do not need a circulatory system because diffusion allows adequate exchange of water, nutrients, and waste, as well as dissolved gases, as shown in [link]**a**. Organisms that are more complex but still only have two layers of cells in their body plan, such as jellies (Cnidaria) and comb jellies (Ctenophora) also use diffusion through their epidermis and internally through the gastrovascular compartment. Both their internal and external tissues are bathed in an aqueous environment and exchange fluids by diffusion on both sides, as illustrated in [link]**b**. Exchange of fluids is assisted by the pulsing of the jellyfish body.

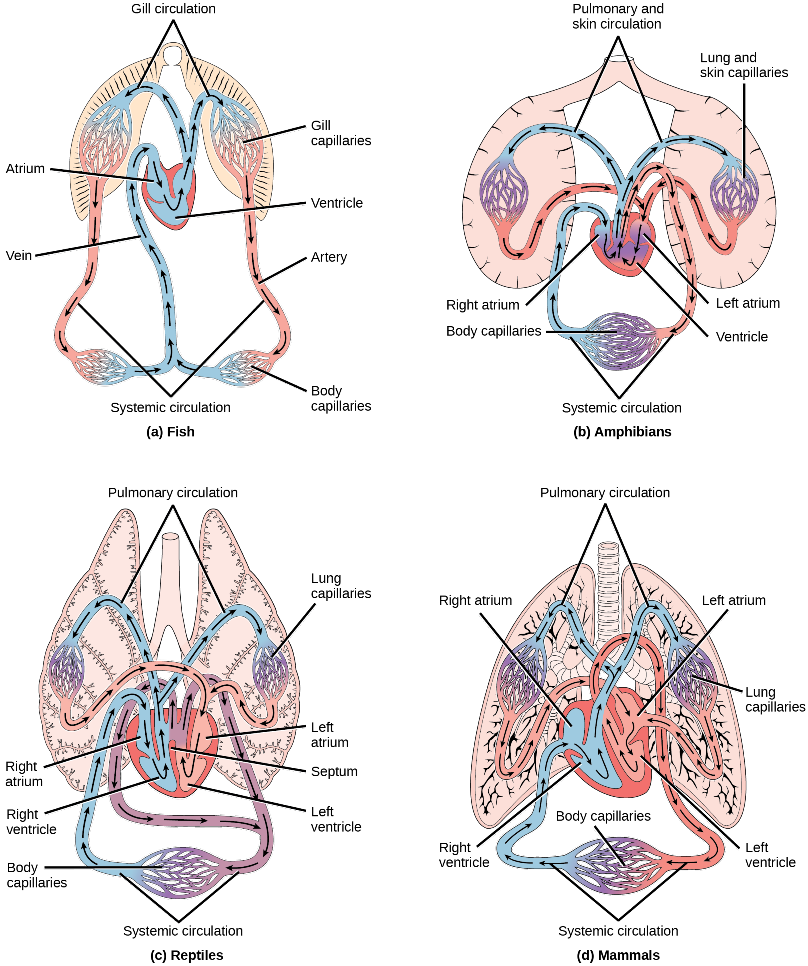


(a) Sponge

(b) Jellyfish

For more complex organisms, diffusion is not efficient for cycling gases, nutrients, and waste effectively through the body; therefore, more complex circulatory systems evolved. Most arthropods and many mollusks have open circulatory systems. In an open system, an elongated beating heart pushes the hemolymph through the body and muscle contractions help to move fluids. The larger more complex crustaceans, including lobsters, have developed arterial-like vessels to push blood through their bodies, and the most active mollusks, such as squids, have evolved a closed circulatory system and are able to move rapidly to catch prey. Closed circulatory systems are a characteristic of vertebrates; however, there are significant differences in the structure of the heart and the circulation of blood between the different vertebrate groups due to adaptation during evolution and associated differences in anatomy. [link] illustrates the basic circulatory systems of some vertebrates: fish, amphibians, reptiles, and

mammals.



As illustrated in [link]**a**. Fish have a single circuit for blood flow and a two-chambered heart that has only a single atrium and a single ventricle. The atrium collects blood that has returned from the body and the ventricle pumps the blood to the gills where gas exchange occurs and the blood is re-oxygenated; this is called **gill circulation**. The blood then continues through the rest of the body

before arriving back at the atrium; this is called **systemic circulation**. This unidirectional flow of blood produces a gradient of oxygenated to deoxygenated blood around the fish's systemic circuit. The result is a limit in the amount of oxygen that can reach some of the organs and tissues of the body, reducing the overall metabolic capacity of fish.

In amphibians, reptiles, birds, and mammals, blood flow is directed in two circuits: one through the lungs and back to the heart, which is called **pulmonary circulation**, and the other throughout the rest of the body and its organs including the brain (systemic circulation). In amphibians, gas exchange also occurs through the skin during pulmonary circulation and is referred to as **pulmocutaneous circulation**.

As shown in [link]**b**, amphibians have a three-chambered heart that has two atria and one ventricle rather than the two-chambered heart of fish. The two **atria** (superior heart chambers) receive blood from the two different circuits (the lungs and the systems), and then there is some mixing of the blood in the heart's **ventricle** (inferior heart chamber), which reduces the efficiency of oxygenation. The advantage to this arrangement is that high pressure in the vessels pushes blood to the lungs and body. The mixing is mitigated by a ridge within the ventricle that diverts oxygen-rich blood

through the systemic circulatory system and deoxygenated blood to the pulmocutaneous circuit. For this reason, amphibians are often described as having **double circulation**.

Most reptiles also have a three-chambered heart similar to the amphibian heart that directs blood to the pulmonary and systemic circuits, as shown in [link]**c**. The ventricle is divided more effectively by a partial septum, which results in less mixing of oxygenated and deoxygenated blood. Some reptiles (alligators and crocodiles) are the most primitive animals to exhibit a four-chambered heart. Crocodilians have a unique circulatory mechanism where the heart shunts blood from the lungs toward the stomach and other organs during long periods of submergence, for instance, while the animal waits for prey or stays underwater waiting for prey to rot. One adaptation includes two main arteries that leave the same part of the heart: one takes blood to the lungs and the other provides an alternate route to the stomach and other parts of the body. Two other adaptations include a hole in the heart between the two ventricles, called the foramen of Panizza, which allows blood to move from one side of the heart to the other, and specialized connective tissue that slows the blood flow to the lungs. Together these adaptations have made crocodiles and alligators one of the most evolutionarily successful animal groups on earth.

In mammals and birds, the heart is also divided into four chambers: two atria and two ventricles, as illustrated in [link]**d**. The oxygenated blood is separated from the deoxygenated blood, which improves the efficiency of double circulation and is probably required for the warm-blooded lifestyle of mammals and birds. The four-chambered heart of birds and mammals evolved independently from a three-chambered heart. The independent evolution of the same or a similar biological trait is referred to as convergent evolution.

Section Summary

In most animals, the circulatory system is used to transport blood through the body. Some primitive animals use diffusion for the exchange of water, nutrients, and gases. However, complex organisms use the circulatory system to carry gases, nutrients, and waste through the body. Circulatory systems may be open (mixed with the interstitial fluid) or closed (separated from the interstitial fluid). Closed circulatory systems are a characteristic of vertebrates; however, there are significant differences in the structure of the heart and the circulation of blood between the different vertebrate groups due to adaptions during evolution and associated differences in anatomy. Fish have a two-chambered heart with unidirectional circulation. Amphibians have a three-chambered heart, which

has some mixing of the blood, and they have double circulation. Most non-avian reptiles have a three-chambered heart, but have little mixing of the blood; they have double circulation. Mammals and birds have a four-chambered heart with no mixing of the blood and double circulation.

Review Questions

Why are open circulatory systems advantageous to some animals?

1. They use less metabolic energy.
2. They help the animal move faster.
3. They do not need a heart.
4. They help large insects develop.

A

Some animals use diffusion instead of a circulatory system. Examples include:

1. birds and jellyfish
2. flatworms and arthropods
3. mollusks and jellyfish
4. none of the above

D

Blood flow that is directed through the lungs and back to the heart is called ______.

1. unidirectional circulation
2. gill circulation
3. pulmonary circulation
4. pulmocutaneous circulation

C

Critical Thinking Questions

Describe a closed circulatory system.

A closed circulatory system is a closed-loop system, in which blood is not free in a cavity. Blood is separate from the bodily interstitial fluid and contained within blood vessels. In this type of system, blood circulates unidirectionally from the heart around the systemic circulatory route, and then returns to the heart.

Describe systemic circulation.

Systemic circulation flows through the systems of the body. The blood flows away from the heart to the brain, liver, kidneys, stomach, and other organs, the limbs, and the muscles of the body; it then returns to the heart.

Glossary

atrium
: (plural: atria) chamber of the heart that receives blood from the veins and sends blood to the ventricles

closed circulatory system
: system in which the blood is separated from the bodily interstitial fluid and contained in blood vessels

double circulation
: flow of blood in two circuits: the pulmonary circuit through the lungs and the systemic circuit through the organs and body

gill circulation
: circulatory system that is specific to animals with gills for gas exchange; the blood flows through the gills for oxygenation

hemocoel
: cavity into which blood is pumped in an open circulatory system

hemolymph
: mixture of blood and interstitial fluid that is found in insects and other arthropods as well as most mollusks

interstitial fluid
: fluid between cells

open circulatory system
: system in which the blood is mixed with interstitial fluid and directly covers the organs

ostium
: (plural: ostia) holes between blood vessels that allow the movement of hemolymph through the body of insects, arthropods, and mollusks with open circulatory systems

pulmocutaneous circulation
: circulatory system in amphibians; the flow of blood to the lungs and the moist skin for gas exchange

pulmonary circulation
: flow of blood away from the heart through the lungs where oxygenation occurs and then returns to the heart again

systemic circulation
: flow of blood away from the heart to the brain, liver, kidneys, stomach, and other organs, the limbs, and the muscles of the body, and then the return of this blood to the heart

unidirectional circulation
: flow of blood in a single circuit; occurs in fish where the blood flows through the gills, then past the organs and the rest of the body, before returning to the heart

ventricle
: (heart) large inferior chamber of the heart that pumps blood into arteries

Components of the Blood

By the end of this section, you will be able to do the following:

- List the basic components of the blood
- Compare red and white blood cells
- Describe blood plasma and serum

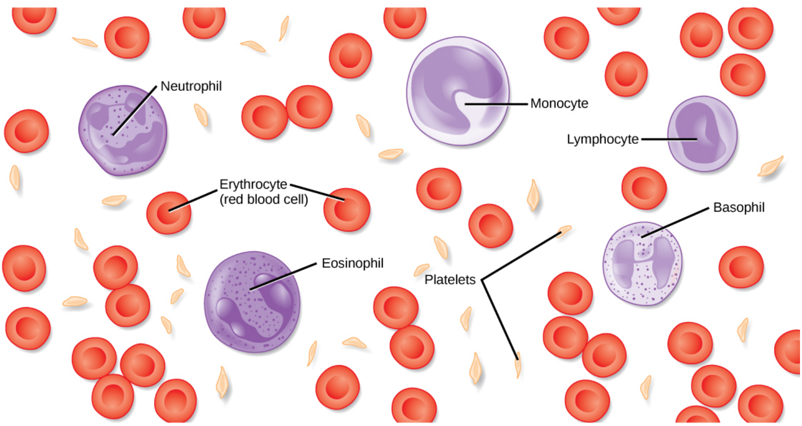
Hemoglobin is responsible for distributing oxygen, and to a lesser extent, carbon dioxide, throughout the circulatory systems of humans, vertebrates, and many invertebrates. The blood is more than the proteins, though. Blood is actually a term used to describe the liquid that moves through the vessels and includes **plasma** (the liquid portion, which contains water, proteins, salts, lipids, and glucose) and the cells (red and white cells) and cell fragments called **platelets**. Blood plasma is actually the dominant component of blood and contains the water, proteins, electrolytes, lipids, and glucose. The cells are responsible for carrying the gases (red cells) and the immune response (white). The platelets are responsible for blood clotting. Interstitial fluid that surrounds cells is separate from the blood, but in hemolymph, they are combined. In humans, cellular components make up approximately 45 percent of the blood and the liquid plasma 55 percent. Blood is 20 percent of a person’s extracellular fluid and eight percent of weight.

The cells and cellular components of human blood

are shown. Red blood cells deliver oxygen to the cells and remove carbon dioxide. White blood cells —including neutrophils, monocytes, lymphocytes, eosinophils, and basophils—are involved in the immune response. Platelets form clots that prevent blood loss after injury.

The Role of Blood in the Body

Blood, like the human blood illustrated in [link] is important for regulation of the body's systems and homeostasis. Blood helps maintain homeostasis by stabilizing pH, temperature, osmotic pressure, and by eliminating excess heat. Blood supports growth by distributing nutrients and hormones, and by removing waste. Blood plays a protective role by transporting clotting factors and platelets to prevent blood loss and transporting the disease-fighting agents or **white blood cells** to sites of infection.



In most vertebrates, (a) hemoglobin delivers oxygen to the body and removes some carbon dioxide.

Hemoglobin is composed of four protein subunits, two alpha chains and two beta chains, and a heme group that has iron associated with it. The iron reversibly associates with oxygen, and in so doing is oxidized from Fe^{2+} to Fe^{3+}. In most mollusks and some arthropods, (b) hemocyanin delivers oxygen. Unlike hemoglobin, hemolymph is not carried in blood cells, but floats free in the hemolymph. Copper instead of iron binds the oxygen, giving the hemolymph a blue-green color. In annelids, such as the earthworm, and some other invertebrates, (c) hemerythrin carries oxygen. Like hemoglobin, hemerythrin is carried in blood cells and has iron associated with it, but despite its name, hemerythrin does not contain heme.

Red Blood Cells

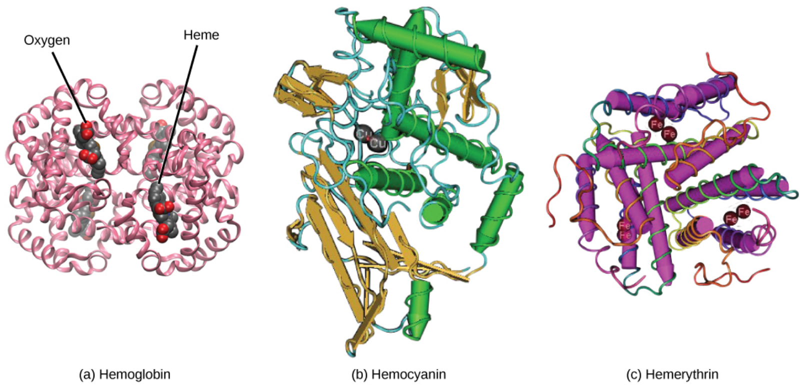
Red blood cells, or erythrocytes (erythro- = “red”; -cyte = “cell”), are specialized cells that circulate through the body delivering oxygen to cells; they are formed from stem cells in the bone marrow. In mammals, red blood cells are small biconcave cells that at maturity do not contain a nucleus or mitochondria and are only 7–8 µm in size. In birds and non-avian reptiles, a nucleus is still maintained in red blood cells.

The red coloring of blood comes from the iron-containing protein hemoglobin, illustrated in [link]**a**. The principle job of this protein is to carry

oxygen, but it also transports carbon dioxide as well. Hemoglobin is packed into red blood cells at a rate of about 250 million molecules of hemoglobin per cell. Each hemoglobin molecule binds four oxygen molecules so that each red blood cell carries one billion molecules of oxygen. There are approximately 25 trillion red blood cells in the five liters of blood in the human body, which could carry up to 25 sextillion (25×10^{21}) molecules of oxygen in the body at any time. In mammals, the lack of organelles in erythrocytes leaves more room for the hemoglobin molecules, and the lack of mitochondria also prevents use of the oxygen for metabolic respiration. Only mammals have anucleated red blood cells, and some mammals (camels, for instance) even have nucleated red blood cells. The advantage of nucleated red blood cells is that these cells can undergo mitosis. Anucleated red blood cells metabolize anaerobically (without oxygen), making use of a primitive metabolic pathway to produce ATP and increase the efficiency of oxygen transport.

Not all organisms use hemoglobin as the method of oxygen transport. Invertebrates that utilize hemolymph rather than blood use different pigments to bind to the oxygen. These pigments use copper or iron to the oxygen. Invertebrates have a variety of other respiratory pigments. Hemocyanin, a blue-green, copper-containing protein, illustrated in [link]**b** is found in mollusks, crustaceans, and

some of the arthropods. Chlorocruorin, a green-colored, iron-containing pigment is found in four families of polychaete tubeworms. Hemerythrin, a red, iron-containing protein is found in some polychaete worms and annelids and is illustrated in [link]**c**. Despite the name, hemerythrin does not contain a heme group and its oxygen-carrying capacity is poor compared to hemoglobin.



The small size and large surface area of red blood cells allows for rapid diffusion of oxygen and carbon dioxide across the plasma membrane. In the lungs, carbon dioxide is released and oxygen is taken in by the blood. In the tissues, oxygen is released from the blood and carbon dioxide is bound for transport back to the lungs. Studies have found that hemoglobin also binds nitrous oxide (NO). NO is a vasodilator that relaxes the blood vessels and capillaries and may help with gas exchange and the passage of red blood cells through narrow vessels. Nitroglycerin, a heart medication for angina and heart attacks, is converted to NO to help relax the blood vessels and increase oxygen flow through the

body.

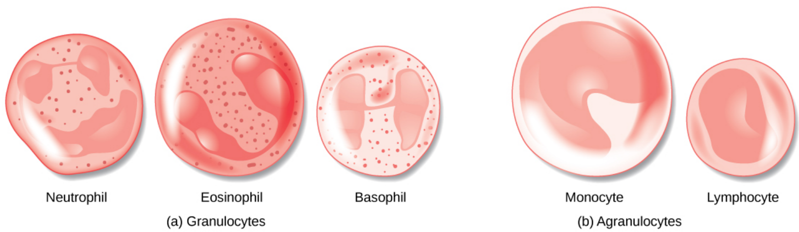
A characteristic of red blood cells is their glycolipid and glycoprotein coating; these are lipids and proteins that have carbohydrate molecules attached. In humans, the surface glycoproteins and glycolipids on red blood cells vary between individuals, producing the different blood types, such as A, B, and O. Red blood cells have an average life span of 120 days, at which time they are broken down and recycled in the liver and spleen by phagocytic macrophages, a type of white blood cell.
(a) Granulocytes—including neutrophils, eosinophils and basophils—are characterized by a lobed nucleus and granular inclusions in the cytoplasm. Granulocytes are typically first-responders during injury or infection. (b) Agranulocytes include lymphocytes and monocytes. Lymphocytes, including B and T cells, are responsible for adaptive immune response. Monocytes differentiate into macrophages and dendritic cells, which in turn respond to infection or injury.

White Blood Cells

White blood cells, also called leukocytes (leuko = white), make up approximately one percent by volume of the cells in blood. The role of white blood cells is very different than that of red blood cells: they are primarily involved in the immune response

to identify and target pathogens, such as invading bacteria, viruses, and other foreign organisms. White blood cells are formed continually; some only live for hours or days, but some live for years.

The morphology of white blood cells differs significantly from red blood cells. They have nuclei and do not contain hemoglobin. The different types of white blood cells are identified by their microscopic appearance after histologic staining, and each has a different specialized function. The two main groups, both illustrated in [link] are the granulocytes, which include the neutrophils, eosinophils, and basophils, and the agranulocytes, which include the monocytes and lymphocytes.



Granulocytes contain granules in their cytoplasm; the agranulocytes are so named because of the lack of granules in their cytoplasm. Some leukocytes become macrophages that either stay at the same site or move through the bloodstream and gather at sites of infection or inflammation where they are attracted by chemical signals from foreign particles and damaged cells. Lymphocytes are the primary cells of the immune system and include B cells, T cells, and natural killer cells. B cells destroy bacteria

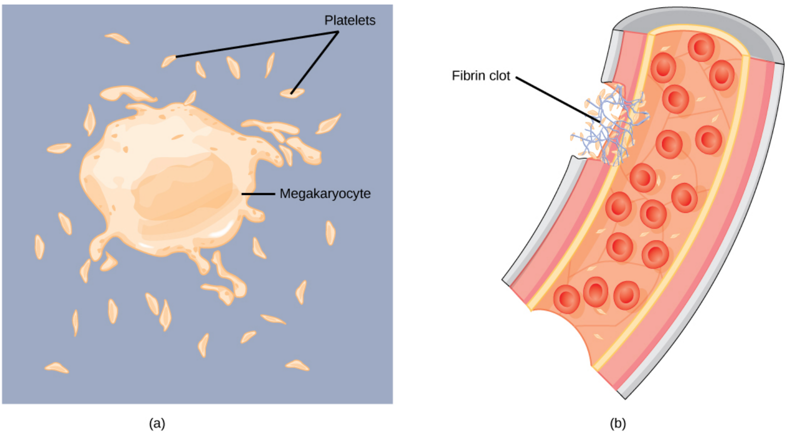
and inactivate their toxins. They also produce antibodies. T cells attack viruses, fungi, some bacteria, transplanted cells, and cancer cells. T cells attack viruses by releasing toxins that kill the viruses. Natural killer cells attack a variety of infectious microbes and certain tumor cells.

One reason that HIV poses significant management challenges is because the virus directly targets T cells by gaining entry through a receptor. Once inside the cell, HIV then multiplies using the T cell's own genetic machinery. After the HIV virus replicates, it is transmitted directly from the infected T cell to macrophages. The presence of HIV can remain unrecognized for an extensive period of time before full disease symptoms develop.
(a) Platelets are formed from large cells called megakaryocytes. The megakaryocyte breaks up into thousands of fragments that become platelets. (b) Platelets are required for clotting of the blood. The platelets collect at a wound site in conjunction with other clotting factors, such as fibrinogen, to form a fibrin clot that prevents blood loss and allows the wound to heal.

Platelets and Coagulation Factors

Blood must clot to heal wounds and prevent excess blood loss. Small cell fragments called platelets (thrombocytes) are attracted to the wound site where they adhere by extending many projections

and releasing their contents. These contents activate other platelets and also interact with other coagulation factors, which convert fibrinogen, a water-soluble protein present in blood serum into fibrin (a non-water soluble protein), causing the blood to clot. Many of the clotting factors require vitamin K to work, and vitamin K deficiency can lead to problems with blood clotting. Many platelets converge and stick together at the wound site forming a platelet plug (also called a fibrin clot), as illustrated in [link]**b**. The plug or clot lasts for a number of days and stops the loss of blood. Platelets are formed from the disintegration of larger cells called megakaryocytes, like that shown in [link]**a**. For each megakaryocyte, 2000–3000 platelets are formed with 150,000 to 400,000 platelets present in each cubic millimeter of blood. Each platelet is disc shaped and 2–4 μm in diameter. They contain many small vesicles but do not contain a nucleus.



Plasma and Serum

The liquid component of blood is called plasma, and it is separated by spinning or centrifuging the blood at high rotations (3000 rpm or higher). The blood cells and platelets are separated by centrifugal forces to the bottom of a specimen tube. The upper liquid layer, the plasma, consists of 90 percent water along with various substances required for maintaining the body's pH, osmotic load, and for protecting the body. The plasma also contains the coagulation factors and antibodies.

The plasma component of blood without the coagulation factors is called the **serum**. Serum is similar to interstitial fluid in which the correct composition of key ions acting as electrolytes is essential for normal functioning of muscles and nerves. Other components in the serum include proteins that assist with maintaining pH and osmotic balance while giving viscosity to the blood. The serum also contains antibodies, specialized proteins that are important for defense against viruses and bacteria. Lipids, including cholesterol, are also transported in the serum, along with various other substances including nutrients, hormones, metabolic waste, plus external substances, such as, drugs, viruses, and bacteria.

Human serum albumin is the most abundant protein in human blood plasma and is synthesized in the

liver. Albumin, which constitutes about half of the blood serum protein, transports hormones and fatty acids, buffers pH, and maintains osmotic pressures. Immunoglobin is a protein antibody produced in the mucosal lining and plays an important role in antibody mediated immunity.

Evolution Connection

Blood Types Related to Proteins on the Surface of the Red Blood Cells

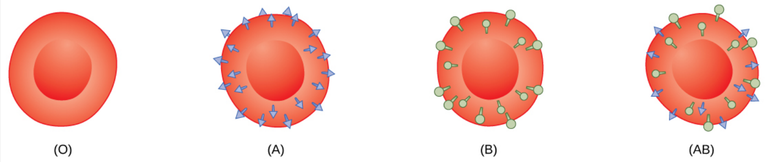
Red blood cells are coated in antigens made of glycolipids and glycoproteins. The composition of these molecules is determined by genetics, which have evolved over time. In humans, the different surface antigens are grouped into 24 different blood groups with more than 100 different antigens on each red blood cell. The two most well known blood groups are the ABO, shown in [link], and Rh systems. The surface antigens in the ABO blood group are glycolipids, called antigen A and antigen B. People with blood type A have antigen A, those with blood type B have antigen B, those with blood type AB have both antigens, and people with blood type O have neither antigen. Antibodies called agglutinougens are found in the blood plasma and react with the A or B antigens, if the two are mixed. When type A and type B blood are combined, agglutination (clumping) of the blood occurs because of antibodies in the plasma that

bind with the opposing antigen; this causes clots that coagulate in the kidney causing kidney failure. Type O blood has neither A or B antigens, and therefore, type O blood can be given to all blood types. Type O negative blood is the universal donor. Type AB positive blood is the universal acceptor because it has both A and B antigen. The ABO blood groups were discovered in 1900 and 1901 by Karl Landsteiner at the University of Vienna.

The Rh blood group was first discovered in Rhesus monkeys. Most people have the Rh antigen (Rh+) and do not have anti-Rh antibodies in their blood. The few people who do not have the Rh antigen and are Rh– can develop anti-Rh antibodies if exposed to Rh+ blood. This can happen after a blood transfusion or after an Rh– woman has an Rh+ baby. The first exposure does not usually cause a reaction; however, at the second exposure, enough antibodies have built up in the blood to produce a reaction that causes agglutination and breakdown of red blood cells. An injection can prevent this reaction.

Human red blood cells may have either type A or B glycoproteins on their surface, both glycoproteins combined (AB), or neither (O). The glycoproteins serve as antigens and can elicit an immune response in a person who receives a transfusion containing unfamiliar antigens. Type O blood, which has no A or B antigens, does not elicit an immune response when injected into a person of

any blood type. Thus, O is considered the universal donor. Persons with type AB blood can accept blood from any blood type, and type AB is considered the universal acceptor.



Link to Learning

Play a blood typing game on the Nobel Prize website to solidify your understanding of blood types.

Section Summary

Specific components of the blood include red blood cells, white blood cells, platelets, and the plasma, which contains coagulation factors and serum. Blood is important for regulation of the body’s pH, temperature, osmotic pressure, the circulation of nutrients and removal of waste, the distribution of hormones from endocrine glands, and the elimination of excess heat; it also contains components for blood clotting. Red blood cells are

specialized cells that contain hemoglobin and circulate through the body delivering oxygen to cells. White blood cells are involved in the immune response to identify and target invading bacteria, viruses, and other foreign organisms; they also recycle waste components, such as old red blood cells. Platelets and blood clotting factors cause the change of the soluble protein fibrinogen to the insoluble protein fibrin at a wound site forming a plug. Plasma consists of 90 percent water along with various substances, such as coagulation factors and antibodies. The serum is the plasma component of the blood without the coagulation factors.

Review Questions

White blood cells:

1. can be classified as granulocytes or agranulocytes
2. defend the body against bacteria and viruses
3. are also called leucocytes
4. all of the above

D

Platelet plug formation occurs at which point?

1. when large megakaryocytes break up into thousands of smaller fragments
2. when platelets are dispersed through the bloodstream
3. when platelets are attracted to a site of blood vessel damage
4. none of the above

C

In humans, the plasma comprises what percentage of the blood?

1. 45 percent
2. 55 percent
3. 25 percent
4. 90 percent

B

The red blood cells of birds differ from mammalian red blood cells because:

1. they are white and have nuclei
2. they do not have nuclei
3. they have nuclei

4. they fight disease

C

Critical Thinking Questions

Describe the cause of different blood type groups.

Red blood cells are coated with proteins called antigens made of glycolipids and glycoproteins. When type A and type B blood are mixed, the blood agglutinates because of antibodies in the plasma that bind with the opposing antigen. Type O blood has no antigens. The Rh blood group has either the Rh antigen (Rh+) or no Rh antigen (Rh–).

List some of the functions of blood in the body.

Blood is important for regulation of the body's pH, temperature, and osmotic pressure, the circulation of nutrients and removal of wastes, the distribution of hormones from endocrine glands, the elimination of excess heat; it also

contains components for the clotting of blood to prevent blood loss. Blood also transports clotting factors and disease-fighting agents.

How does the lymphatic system work with blood flow?

Lymph capillaries take fluid from the blood to the lymph nodes. The lymph nodes filter the lymph by percolation through connective tissue filled with white blood cells. The white blood cells remove infectious agents, such as bacteria and viruses, to clean the lymph before it returns to the bloodstream.

Glossary

plasma
: liquid component of blood that is left after the cells are removed

platelet
: (also, thrombocyte) small cellular fragment that collects at wounds, cross-reacts with clotting factors, and forms a plug to prevent blood loss

red blood cell
: small (7–8 μm) biconcave cell without

mitochondria (and in mammals without nuclei) that is packed with hemoglobin, giving the cell its red color; transports oxygen through the body

serum
: plasma without the coagulation factors

white blood cell
: large (30 μm) cell with nuclei of which there are many types with different roles including the protection of the body from viruses and bacteria, and cleaning up dead cells and other waste

Mammalian Heart and Blood Vessels

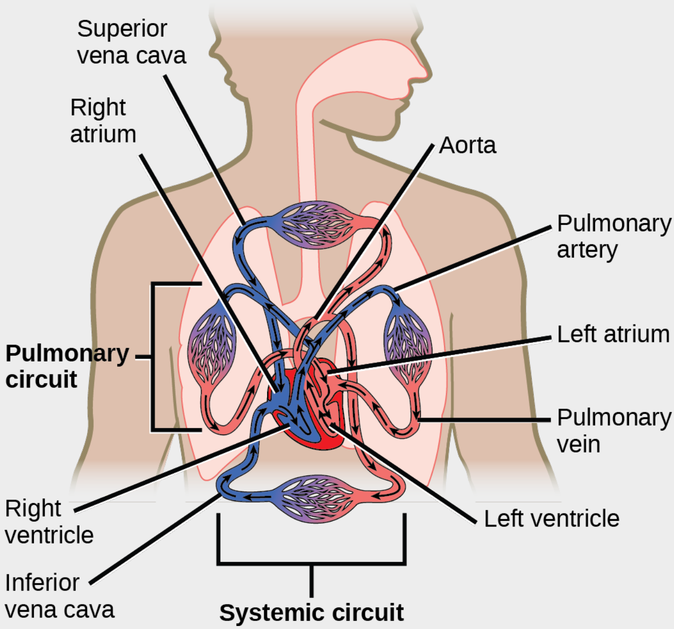
By the end of this section, you will be able to do the following:

- Describe the structure of the heart and explain how cardiac muscle is different from other muscles
- Describe the cardiac cycle
- Explain the structure of arteries, veins, and capillaries, and how blood flows through the body

The heart is a complex muscle that pumps blood through the three divisions of the circulatory system: the coronary (vessels that serve the heart), pulmonary (heart and lungs), and systemic (systems of the body), as shown in [link]. Coronary circulation intrinsic to the heart takes blood directly from the main artery (aorta) coming from the heart. For pulmonary and systemic circulation, the heart has to pump blood to the lungs or the rest of the body, respectively. In vertebrates, the lungs are relatively close to the heart in the thoracic cavity. The shorter distance to pump means that the muscle wall on the right side of the heart is not as thick as the left side which must have enough pressure to pump blood all the way to your big toe.

Visual Connection

The mammalian circulatory system is divided into three circuits: the systemic circuit, the pulmonary circuit, and the coronary circuit. Blood is pumped from veins of the systemic circuit into the right atrium of the heart, then into the right ventricle. Blood then enters the pulmonary circuit, and is oxygenated by the lungs. From the pulmonary circuit, blood reenters the heart through the left atrium. From the left ventricle, blood reenters the systemic circuit through the aorta and is distributed to the rest of the body. The coronary circuit, which provides blood to the heart, is not shown.



Which of the following statements about the circulatory system is false?

1. Blood in the pulmonary vein is deoxygenated.

2. Blood in the inferior vena cava is deoxygenated.
3. Blood in the pulmonary artery is deoxygenated.
4. Blood in the aorta is oxygenated.

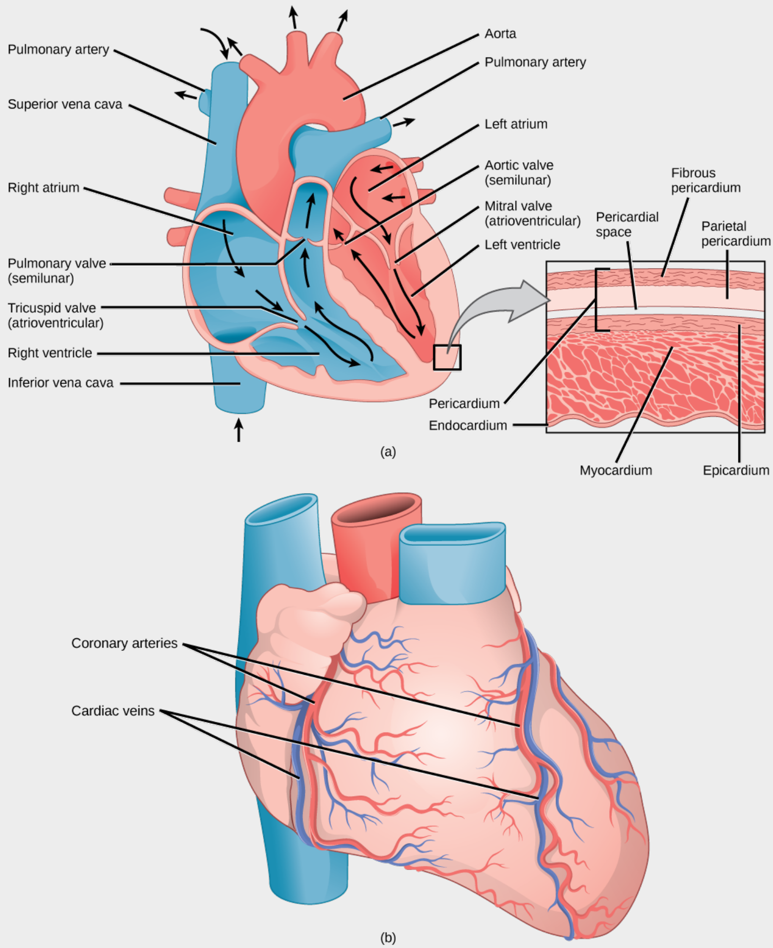
Structure of the Heart

The heart muscle is asymmetrical as a result of the distance blood must travel in the pulmonary and systemic circuits. Since the right side of the heart sends blood to the pulmonary circuit it is smaller than the left side which must send blood out to the whole body in the systemic circuit, as shown in [link]. In humans, the heart is about the size of a clenched fist; it is divided into four chambers: two atria and two ventricles. There is one atrium and one ventricle on the right side and one atrium and one ventricle on the left side. The atria are the chambers that receive blood, and the ventricles are the chambers that pump blood. The right atrium receives deoxygenated blood from the **superior vena cava**, which drains blood from the jugular vein that comes from the brain and from the veins that come from the arms, as well as from the **inferior vena cava** which drains blood from the veins that come from the lower organs and the legs.

In addition, the right atrium receives blood from the coronary sinus which drains deoxygenated blood from the heart itself. This deoxygenated blood then passes to the right ventricle through the **atrioventricular valve** or the **tricuspid valve**, a flap of connective tissue that opens in only one direction to prevent the backflow of blood. The valve separating the chambers on the left side of the heart valve is called the biscuspid or mitral valve. After it is filled, the right ventricle pumps the blood through the pulmonary arteries, bypassing the **semilunar valve** (or pulmonic valve) to the lungs for re-oxygenation. After blood passes through the pulmonary arteries, the right semilunar valves close preventing the blood from flowing backwards into the right ventricle. The left atrium then receives the oxygen-rich blood from the lungs via the pulmonary veins. This blood passes through the **bicuspid valve** or mitral valve (the atrioventricular valve on the left side of the heart) to the left ventricle where the blood is pumped out through the **aorta**, the major artery of the body, taking oxygenated blood to the organs and muscles of the body. Once blood is pumped out of the left ventricle and into the aorta, the aortic semilunar valve (or aortic valve) closes preventing blood from flowing backward into the left ventricle. This pattern of pumping is referred to as double circulation and is found in all mammals.

Visual Connection

(a) The heart is primarily made of a thick muscle layer, called the myocardium, surrounded by membranes. One-way valves separate the four chambers. (b) Blood vessels of the coronary system, including the coronary arteries and veins, keep the heart musculature oxygenated.



Which of the following statements about the heart is false?

1. The mitral valve separates the left ventricle from the left atrium.
2. Blood travels through the bicuspid valve to the left atrium.
3. Both the aortic and the pulmonary valves are semilunar valves.
4. The mitral valve is an atrioventricular valve.

The heart is composed of three layers; the epicardium, the myocardium, and the endocardium, illustrated in [link]. The inner wall of the heart has a lining called the **endocardium**. The **myocardium** consists of the heart muscle cells that make up the middle layer and the bulk of the heart wall. The outer layer of cells is called the **epicardium**, of which the second layer is a membranous layered structure called the **pericardium** that surrounds and protects the heart; it allows enough room for vigorous pumping but also keeps the heart in place to reduce friction between the heart and other structures.

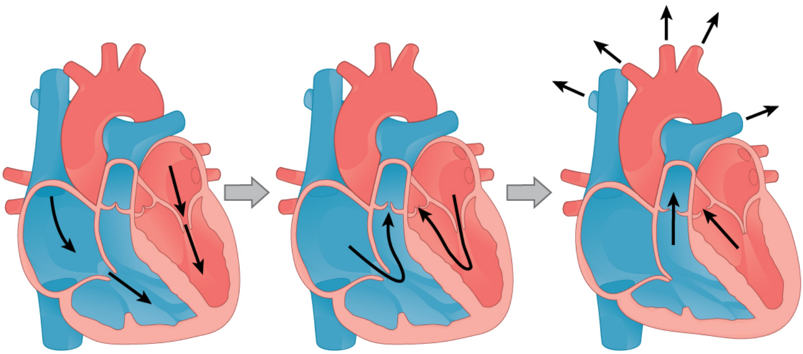
The heart has its own blood vessels that supply the heart muscle with blood. The **coronary arteries** branch from the aorta and surround the outer surface of the heart like a crown. They diverge into capillaries where the heart muscle is supplied with oxygen before converging again into the **coronary veins** to take the deoxygenated blood back to the

right atrium where the blood will be re-oxygenated through the pulmonary circuit. The heart muscle will die without a steady supply of blood. **Atherosclerosis** is the blockage of an artery by the buildup of fatty plaques. Because of the size (narrow) of the coronary arteries and their function in serving the heart itself, atherosclerosis can be deadly in these arteries. The slowdown of blood flow and subsequent oxygen deprivation that results from atherosclerosis causes severe pain, known as **angina**, and complete blockage of the arteries will cause **myocardial infarction**: the death of cardiac muscle tissue, commonly known as a heart attack.

During (a) cardiac diastole, the heart muscle is relaxed and blood flows into the heart. During (b) atrial systole, the atria contract, pushing blood into the ventricles. During (c) atrial diastole, the ventricles contract, forcing blood out of the heart. Cardiomyocytes are striated muscle cells found in cardiac tissue. (credit: modification of work by Dr. S. Girod, Anton Becker; scale-bar data from Matt Russell) The beating of the heart is regulated by an electrical impulse that causes the characteristic reading of an ECG. The signal is initiated at the sinoatrial valve. The signal then (a) spreads to the atria, causing them to contract. The signal is (b) delayed at the atrioventricular node before it is passed on to the (c) heart apex. The delay allows the atria to relax before the (d) ventricles contract. The final part of the ECG cycle prepares the heart for the next beat.

The Cardiac Cycle

The main purpose of the heart is to pump blood through the body; it does so in a repeating sequence called the cardiac cycle. The **cardiac cycle** is the coordination of the filling and emptying of the heart of blood by electrical signals that cause the heart muscles to contract and relax. The human heart beats over 100,000 times per day. In each cardiac cycle, the heart contracts (**systole**), pushing out the blood and pumping it through the body; this is followed by a relaxation phase (**diastole**), where the heart fills with blood, as illustrated in [link]. The atria contract at the same time, forcing blood through the atrioventricular valves into the ventricles. Closing of the atrioventricular valves produces a monosyllabic “lup” sound. Following a brief delay, the ventricles contract at the same time forcing blood through the semilunar valves into the aorta and the artery transporting blood to the lungs (via the pulmonary artery). Closing of the semilunar valves produces a monosyllabic “dup” sound.

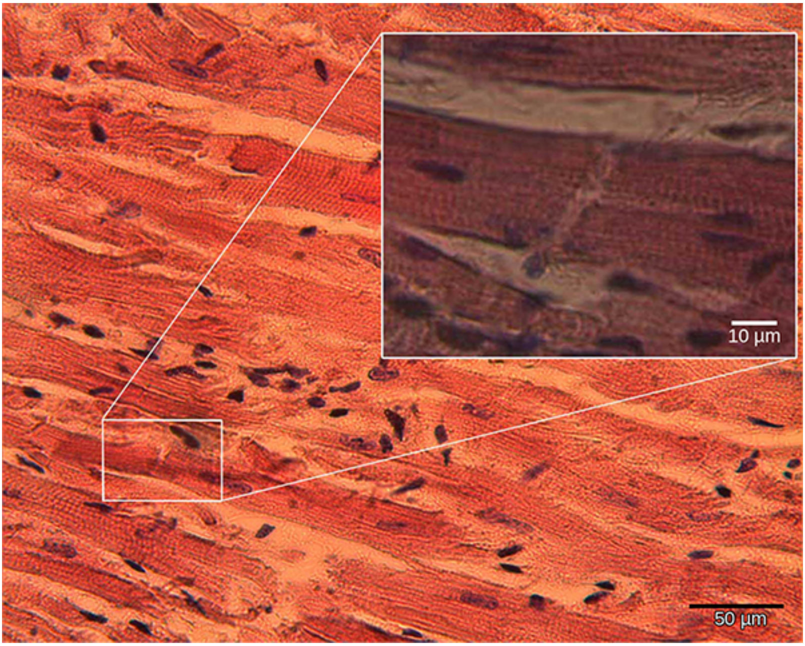


(a) Cardiac diastole: all chambers are relaxed, and blood flows into the heart.

(b) Atrial systole, ventricular diastole: atria contract, pushing blood into the ventricles.

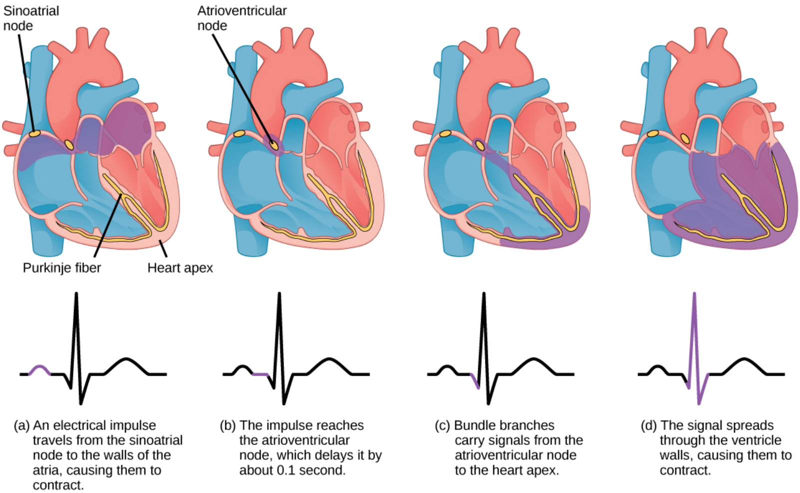
(c) Atrial diastole, ventricular systole: after the atria relax, the ventricles contract, pushing blood out of the heart.

The pumping of the heart is a function of the cardiac muscle cells, or cardiomyocytes, that make up the heart muscle. **Cardiomyocytes**, shown in [link], are distinctive muscle cells that are striated like skeletal muscle but pump rhythmically and involuntarily like smooth muscle; they are connected by intercalated disks exclusive to cardiac muscle. They are self-stimulated for a period of time and isolated cardiomyocytes will beat if given the correct balance of nutrients and electrolytes.



The autonomous beating of cardiac muscle cells is regulated by the heart's internal pacemaker that uses electrical signals to time the beating of the heart. The electrical signals and mechanical actions, illustrated in [link], are intimately intertwined. The internal pacemaker starts at the **sinoatrial (SA) node**, which is located near the wall of the right atrium. Electrical charges spontaneously pulse from the SA node causing the two atria to contract in unison. The pulse reaches a second node, called the atrioventricular (AV) node, between the right atrium and right ventricle where it pauses for approximately 0.1 second before spreading to the walls of the ventricles. From the AV node, the electrical impulse enters the bundle of His, then to the left and right bundle branches extending

through the interventricular septum. Finally, the Purkinje fibers conduct the impulse from the apex of the heart up the ventricular myocardium, and then the ventricles contract. This pause allows the atria to empty completely into the ventricles before the ventricles pump out the blood. The electrical impulses in the heart produce electrical currents that flow through the body and can be measured on the skin using electrodes. This information can be observed as an **electrocardiogram (ECG)**—a recording of the electrical impulses of the cardiac muscle.



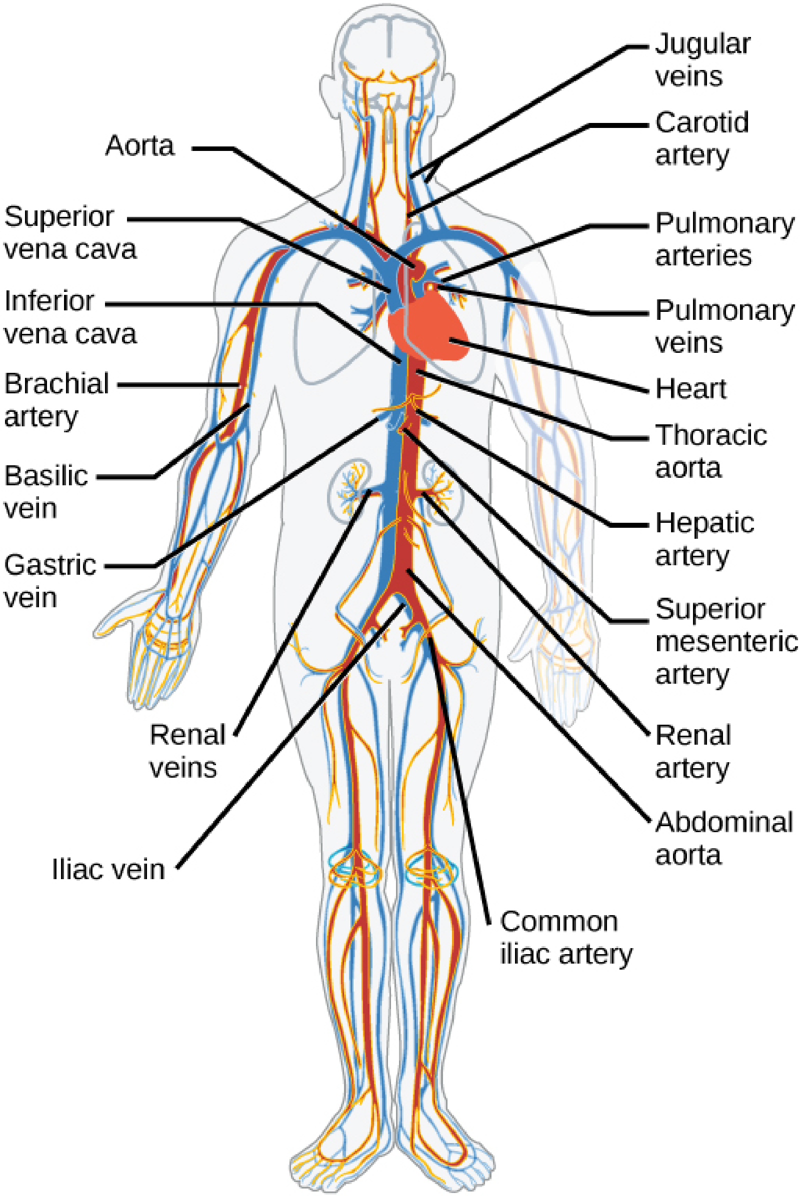
Link to Learning

Visit this site to see the heart's "pacemaker" in action.

The major human arteries and veins are shown. (credit: modification of work by Mariana Ruiz Villareal) Arteries and veins consist of three layers: an outer tunica externa, a middle tunica media, and an inner tunica intima. Capillaries consist of a single layer of epithelial cells, the tunica intima. (credit: modification of work by NCI, NIH)

Arteries, Veins, and Capillaries

The blood from the heart is carried through the body by a complex network of blood vessels ([link]). **Arteries** take blood away from the heart. The main artery is the aorta that branches into major arteries that take blood to different limbs and organs. These major arteries include the carotid artery that takes blood to the brain, the brachial arteries that take blood to the arms, and the thoracic artery that takes blood to the thorax and then into the hepatic, renal, and gastric arteries for the liver, kidney, and stomach, respectively. The iliac artery takes blood to the lower limbs. The major arteries diverge into minor arteries, and then smaller vessels called **arterioles**, to reach more deeply into the muscles and organs of the body.

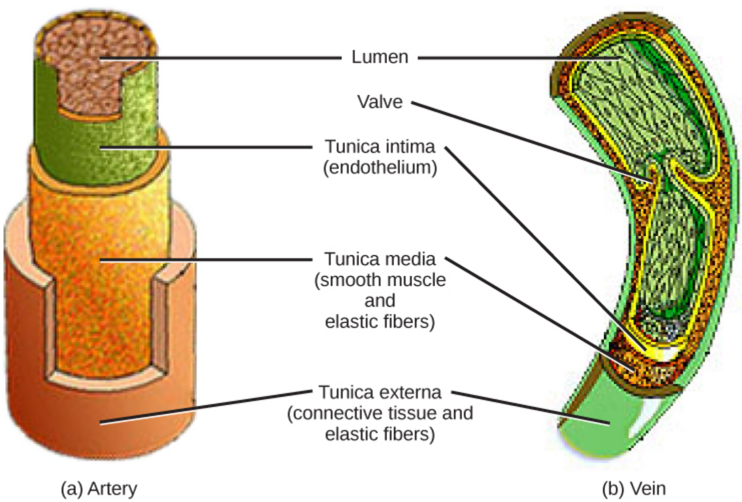


Arterioles diverge into capillary beds. **Capillary beds** contain a large number (10 to 100) of **capillaries** that branch among the cells and tissues

of the body. Capillaries are narrow-diameter tubes that can fit red blood cells through in single file and are the sites for the exchange of nutrients, waste, and oxygen with tissues at the cellular level. Fluid also crosses into the interstitial space from the capillaries. The capillaries converge again into **venules** that connect to minor veins that finally connect to major veins that take blood high in carbon dioxide back to the heart. **Veins** are blood vessels that bring blood back to the heart. The major veins drain blood from the same organs and limbs that the major arteries supply. Fluid is also brought back to the heart via the lymphatic system.

The structure of the different types of blood vessels reflects their function or layers. There are three distinct layers, or tunics, that form the walls of blood vessels ([link]). The first tunic is a smooth, inner lining of endothelial cells that are in contact with the red blood cells. The endothelial tunic is continuous with the endocardium of the heart. In capillaries, this single layer of cells is the location of diffusion of oxygen and carbon dioxide between the endothelial cells and red blood cells, as well as the exchange site via endocytosis and exocytosis. The movement of materials at the site of capillaries is regulated by **vasoconstriction**, narrowing of the blood vessels, and **vasodilation**, widening of the blood vessels; this is important in the overall regulation of blood pressure.

Veins and arteries both have two further tunics that surround the endothelium: the middle tunic is composed of smooth muscle and the outermost layer is connective tissue (collagen and elastic fibers). The elastic connective tissue stretches and supports the blood vessels, and the smooth muscle layer helps regulate blood flow by altering vascular resistance through vasoconstriction and vasodilation. The arteries have thicker smooth muscle and connective tissue than the veins to accommodate the higher pressure and speed of freshly pumped blood. The veins are thinner walled as the pressure and rate of flow are much lower. In addition, veins are structurally different than arteries in that veins have valves to prevent the backflow of blood. Because veins have to work against gravity to get blood back to the heart, contraction of skeletal muscle assists with the flow of blood back to the heart.



Section Summary

The heart muscle pumps blood through three divisions of the circulatory system: coronary, pulmonary, and systemic. There is one atrium and one ventricle on the right side and one atrium and one ventricle on the left side. The pumping of the heart is a function of cardiomyocytes, distinctive muscle cells that are striated like skeletal muscle but pump rhythmically and involuntarily like smooth muscle. The internal pacemaker starts at the sinoatrial node, which is located near the wall of the right atrium. Electrical charges pulse from the SA node causing the two atria to contract in unison; then the pulse reaches the atrioventricular node between the right atrium and right ventricle. A pause in the electric signal allows the atria to empty completely into the ventricles before the ventricles pump out the blood. The blood from the heart is carried through the body by a complex network of blood vessels; arteries take blood away from the heart, and veins bring blood back to the heart.

Visual Connection Questions

[link] Which of the following statements about the circulatory system is false?

1. Blood in the pulmonary vein is deoxygenated.
2. Blood in the inferior vena cava is deoxygenated.
3. Blood in the pulmonary artery is deoxygenated.
4. Blood in the aorta is oxygenated.

[link] C

[link] Which of the following statements about the heart is false?

1. The mitral valve separates the left ventricle from the left atrium.
2. Blood travels through the bicuspid valve to the left atrium.
3. Both the aortic and the pulmonary valves are semilunar valves.
4. The mitral valve is an atrioventricular valve.

[link] B

Review Questions

The heart's internal pacemaker beats by:

1. an internal implant that sends an electrical impulse through the heart
2. the excitation of cardiac muscle cells at the sinoatrial node followed by the atrioventricular node
3. the excitation of cardiac muscle cells at the atrioventricular node followed by the sinoatrial node
4. the action of the sinus

B

During the systolic phase of the cardiac cycle, the heart is ______.

1. contracting
2. relaxing
3. contracting and relaxing
4. filling with blood

A

Cardiomyocytes are similar to skeletal muscle because:

1. they beat involuntarily

2. they are used for weight lifting
3. they pulse rhythmically
4. they are striated

D

How do arteries differ from veins?

1. Arteries have thicker smooth muscle layers to accommodate the changes in pressure from the heart.
2. Arteries carry blood.
3. Arteries have thinner smooth muscle layers and valves and move blood by the action of skeletal muscle.
4. Arteries are thin walled and are used for gas exchange.

A

Critical Thinking Questions

Describe the cardiac cycle.

The heart receives an electrical signal from the

sinoatrial node triggering the cardiac muscle cells in the atria to contract. The signal pauses at the atrioventricular node before spreading to the walls of the ventricles so the blood is pumped through the body. This is the systolic phase. The heart then relaxes in the diastole and fills again with blood.

What happens in capillaries?

The capillaries basically exchange materials with their surroundings. Their walls are very thin and are made of one or two layers of cells, where gases, nutrients, and waste are diffused. They are distributed as beds, complex networks that link arteries as well as veins.

Glossary

angina
: pain caused by partial blockage of the coronary arteries by the buildup of plaque and lack of oxygen to the heart muscle

aorta
: major artery of the body that takes blood away from the heart

arteriole

small vessel that connects an artery to a capillary bed

artery
: blood vessel that takes blood away from the heart

atherosclerosis
: buildup of fatty plaques in the coronary arteries in the heart

atrioventricular valve
: one-way membranous flap of connective tissue between the atrium and the ventricle in the right side of the heart; also known as tricuspid valve

bicuspid valve
: (also, mitral valve; left atrioventricular valve) one-way membranous flap between the atrium and the ventricle in the left side of the heart

capillary
: smallest blood vessel that allows the passage of individual blood cells and the site of diffusion of oxygen and nutrient exchange

capillary bed
: large number of capillaries that converge to take blood to a particular organ or tissue

cardiac cycle
: filling and emptying the heart of blood by electrical signals that cause the heart muscles to contract and relax

cardiomyocyte
: specialized heart muscle cell that is striated but contracts involuntarily like smooth muscle

coronary artery
: vessel that supplies the heart tissue with blood

coronary vein
: vessel that takes blood away from the heart tissue back to the chambers in the heart

diastole
: relaxation phase of the cardiac cycle when the heart is relaxed and the ventricles are filling with blood

electrocardiogram (ECG)
: recording of the electrical impulses of the cardiac muscle

endocardium
: innermost layer of tissue in the heart

epicardium
: outermost tissue layer of the heart

inferior vena cava
: drains blood from the veins that come from the lower organs and the legs

myocardial infarction
: (also, heart attack) complete blockage of the coronary arteries and death of the cardiac muscle tissue

myocardium
: heart muscle cells that make up the middle layer and the bulk of the heart wall

pericardium
: membrane layer protecting the heart; also part of the epicardium

semilunar valve
: membranous flap of connective tissue between the aorta and a ventricle of the heart (the aortic or pulmonary semilunar valves)

sinoatrial (SA) node
: the heart's internal pacemaker; located near the wall of the right atrium

superior vena cava
: drains blood from the jugular vein that comes from the brain and from the veins that come from the arms

systole

contraction phase of cardiac cycle when the ventricles are pumping blood into the arteries

tricuspid valve
: one-way membranous flap of connective tissue between the atrium and the ventricle in the right side of the heart; also known as atrioventricular valve

vasoconstriction
: narrowing of a blood vessel

vasodilation
: widening of a blood vessel

vein
: blood vessel that brings blood back to the heart

vena cava
: major vein of the body returning blood from the upper and lower parts of the body; see the superior vena cava and inferior vena cava

venule
: blood vessel that connects a capillary bed to a vein

Blood Flow and Blood Pressure Regulation
By the end of this section, you will be able to do the following:

- Describe the system of blood flow through the body
- Describe how blood pressure is regulated

Blood pressure (BP) is the pressure exerted by blood on the walls of a blood vessel that helps to push blood through the body. Systolic blood pressure measures the amount of pressure that blood exerts on vessels while the heart is beating. The optimal systolic blood pressure is 120 mmHg. Diastolic blood pressure measures the pressure in the vessels between heartbeats. The optimal diastolic blood pressure is 80 mmHg. Many factors can affect blood pressure, such as hormones, stress, exercise, eating, sitting, and standing. Blood flow through the body is regulated by the size of blood vessels, by the action of smooth muscle, by one-way valves, and by the fluid pressure of the blood itself. Fluid from the capillaries moves into the interstitial space and lymph capillaries by diffusion down a pressure gradient and also by osmosis. Out of 7,200 liters of fluid pumped by the average heart in a day, over 1,500 liters is filtered. (credit: modification of work by NCI, NIH)

How Blood Flows Through the Body

Blood is pushed through the body by the action of the pumping heart. With each rhythmic pump, blood is pushed under high pressure and velocity away from the heart, initially along the main artery, the aorta. In the aorta, the blood travels at 30 cm/sec. As blood moves into the arteries, arterioles, and ultimately to the capillary beds, the rate of movement slows dramatically to about 0.026 cm/sec, one-thousand times slower than the rate of movement in the aorta. While the diameter of each individual arteriole and capillary is far narrower than the diameter of the aorta, and according to the law of continuity, fluid should travel faster through a narrower diameter tube, the rate is actually slower due to the overall diameter of all the combined capillaries being far greater than the diameter of the individual aorta.

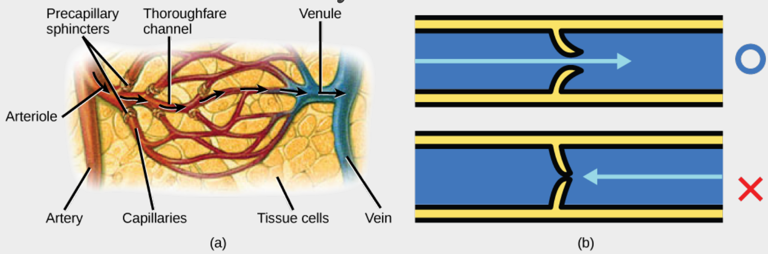
The slow rate of travel through the capillary beds, which reach almost every cell in the body, assists with gas and nutrient exchange and also promotes the diffusion of fluid into the interstitial space. After the blood has passed through the capillary beds to the venules, veins, and finally to the main venae cavae, the rate of flow increases again but is still much slower than the initial rate in the aorta. Blood primarily moves in the veins by the rhythmic movement of smooth muscle in the vessel wall and by the action of the skeletal muscle as the body moves. Because most veins must move blood against the pull of gravity, blood is prevented from flowing

backward in the veins by one-way valves. Because skeletal muscle contraction aids in venous blood flow, it is important to get up and move frequently after long periods of sitting so that blood will not pool in the extremities.

Blood flow through the capillary beds is regulated depending on the body's needs and is directed by nerve and hormone signals. For example, after a large meal, most of the blood is diverted to the stomach by vasodilation of vessels of the digestive system and vasoconstriction of other vessels. During exercise, blood is diverted to the skeletal muscles through vasodilation while blood to the digestive system would be lessened through vasoconstriction. The blood entering some capillary beds is controlled by small muscles, called precapillary sphincters, illustrated in [link]. If the sphincters are open, the blood will flow into the associated branches of the capillary blood. If all of the sphincters are closed, then the blood will flow directly from the arteriole to the venule through the thoroughfare channel (see [link]). These muscles allow the body to precisely control when capillary beds receive blood flow. At any given moment only about 5–10% of our capillary beds actually have blood flowing through them.

Visual Connection

(a) Precapillary sphincters are rings of smooth muscle that regulate the flow of blood through capillaries; they help control the location of blood flow to where it is needed. (b) Valves in the veins prevent blood from moving backward. (credit a: modification of work by NCI)



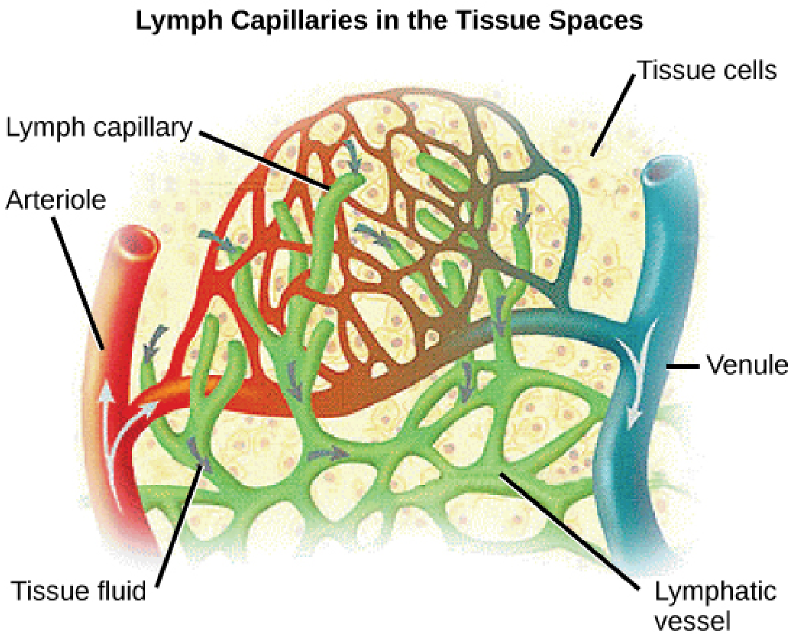
Varicose veins are veins that become enlarged because the valves no longer close properly, allowing blood to flow backward. Varicose veins are often most prominent on the legs. Why do you think this is the case?

Link to Learning

See the circulatory system's blood flow.
https://www.openstax.org/l/circulation

Proteins and other large solutes cannot leave the capillaries. The loss of the watery plasma creates a hyperosmotic solution within the capillaries, especially near the venules. This causes about 85% of the plasma that leaves the capillaries to eventually diffuse back into the capillaries near the

venules. The remaining 15% of blood plasma drains out from the interstitial fluid into nearby lymphatic vessels ([link]). The fluid in the lymph is similar in composition to the interstitial fluid. The lymph fluid passes through lymph nodes before it returns to the heart via the vena cava. **Lymph nodes** are specialized organs that filter the lymph by percolation through a maze of connective tissue filled with white blood cells. The white blood cells remove infectious agents, such as bacteria and viruses, to clean the lymph before it returns to the bloodstream. After it is cleaned, the lymph returns to the heart by the action of smooth muscle pumping, skeletal muscle action, and one-way valves joining the returning blood near the junction of the venae cavae entering the right atrium of the heart.



Evolution Connection

Vertebrate Diversity in Blood Circulation

Blood circulation has evolved differently in vertebrates and may show variation in different animals for the required amount of pressure, organ and vessel location, and organ size. Animals with longs necks and those that live in cold environments have distinct blood pressure adaptations.

Long necked animals, such as giraffes, need to pump blood upward from the heart against gravity. The blood pressure required from the pumping of the left ventricle would be equivalent to 250 mm Hg (mm Hg = millimeters of mercury, a unit of pressure) to reach the height of a giraffe's head, which is 2.5 meters higher than the heart. However, if checks and balances were not in place, this blood pressure would damage the giraffe's brain, particularly if it was bending down to drink. These checks and balances include valves and feedback mechanisms that reduce the rate of cardiac output. Long-necked dinosaurs such as the sauropods had to pump blood even higher, up to ten meters above the heart. This would have required a blood pressure of more than 600 mm Hg, which could only have been achieved by an enormous heart. Evidence for such an enormous heart does not exist and mechanisms to reduce the blood pressure required include the slowing of metabolism as these animals grew larger. It is likely that they did not routinely feed on tree tops

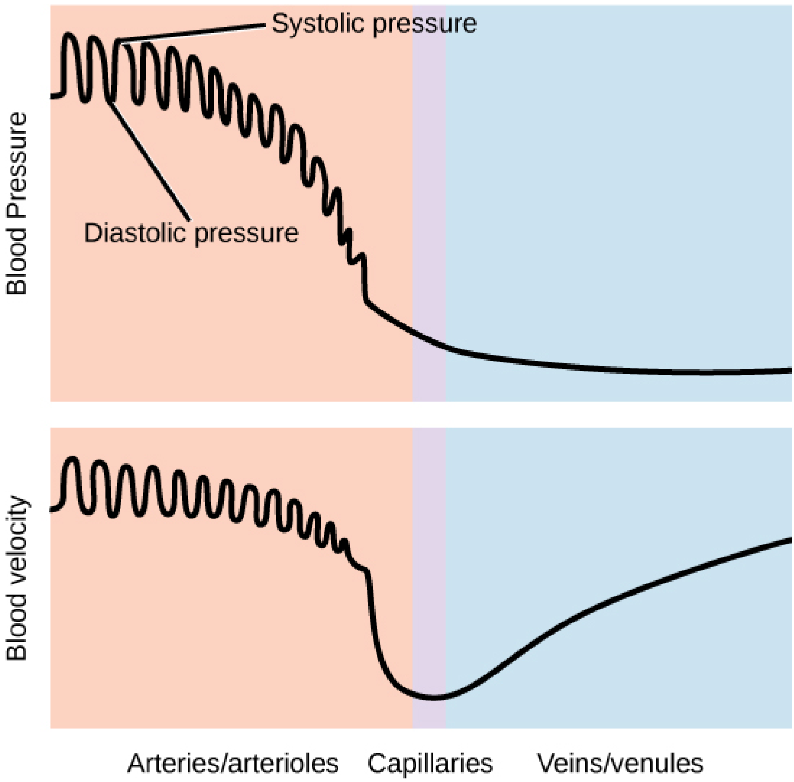
but grazed on the ground.
Living in cold water, whales need to maintain the temperature in their blood. This is achieved by the veins and arteries being close together so that heat exchange can occur. This mechanism is called a countercurrent heat exchanger. The blood vessels and the whole body are also protected by thick layers of blubber to prevent heat loss. In land animals that live in cold environments, thick fur and hibernation are used to retain heat and slow metabolism.

Blood pressure is related to the blood velocity in the arteries and arterioles. In the capillaries and veins, the blood pressure continues to decrease but velocity increases.

Blood Pressure

The pressure of the blood flow in the body is produced by the hydrostatic pressure of the fluid (blood) against the walls of the blood vessels. Fluid will move from areas of high to low hydrostatic pressures. In the arteries, the hydrostatic pressure near the heart is very high and blood flows to the arterioles where the rate of flow is slowed by the narrow openings of the arterioles. During systole, when new blood is entering the arteries, the artery walls stretch to accommodate the increase of

pressure of the extra blood; during diastole, the walls return to normal because of their elastic properties. The blood pressure of the systole phase and the diastole phase, graphed in [link], gives the two pressure readings for blood pressure. For example, 120/80 indicates a reading of 120 mm Hg during the systole and 80 mm Hg during diastole. Throughout the cardiac cycle, the blood continues to empty into the arterioles at a relatively even rate. This resistance to blood flow is called **peripheral resistance**.



Blood Pressure Regulation

Cardiac output is the volume of blood pumped by the heart in one minute. It is calculated by multiplying the number of heart contractions that occur per minute (heart rate) times the **stroke volume** (the volume of blood pumped into the aorta per contraction of the left ventricle). Therefore, cardiac output can be increased by increasing heart rate, as when exercising. However, cardiac output can also be increased by increasing stroke volume, such as if the heart contracts with greater strength. Stroke volume can also be increased by speeding blood circulation through the body so that more blood enters the heart between contractions. During heavy exertion, the blood vessels relax and increase in diameter, offsetting the increased heart rate and ensuring adequate oxygenated blood gets to the muscles. Stress triggers a decrease in the diameter of the blood vessels, consequently increasing blood pressure. These changes can also be caused by nerve signals or hormones, and even standing up or lying down can have a great effect on blood pressure.

Section Summary

Blood primarily moves through the body by the rhythmic movement of smooth muscle in the vessel wall and by the action of the skeletal muscle as the body moves. Blood is prevented from flowing

backward in the veins by one-way valves. Blood flow through the capillary beds is controlled by precapillary sphincters to increase and decrease flow depending on the body’s needs and is directed by nerve and hormone signals. Lymph vessels take fluid that has leaked out of the blood to the lymph nodes where it is cleaned before returning to the heart. During systole, blood enters the arteries, and the artery walls stretch to accommodate the extra blood. During diastole, the artery walls return to normal. The blood pressure of the systole phase and the diastole phase gives the two pressure readings for blood pressure.

Visual Connection Questions

[link] Varicose veins are veins that become enlarged because the valves no longer close properly, allowing blood to flow backward. Varicose veins are often most prominent on the legs. Why do you think this is the case?

[link] Blood in the legs is farthest away from the heart and has to flow up to reach it.

Review Questions

High blood pressure would be a result of ______.

1. a high cardiac output and high peripheral resistance
2. a high cardiac output and low peripheral resistance
3. a low cardiac output and high peripheral resistance
4. a low cardiac output and low peripheral resistance

A

Critical Thinking Questions

How does blood pressure change during heavy exercise?

The heart rate increases, which increases the hydrostatic pressure against the artery walls. At the same time, the arterioles dilate in response to the increased exercise, which reduces peripheral resistance.

Glossary

blood pressure (BP)
: pressure of blood in the arteries that helps to push blood through the body

cardiac output
: the volume of blood pumped by the heart in one minute as a product of heart rate multiplied by stroke volume

lymph node
: specialized organ that contains a large number of macrophages that clean the lymph before the fluid is returned to the heart

peripheral resistance
: resistance of the artery and blood vessel walls to the pressure placed on them by the force of the heart pumping

precapillary sphincter
: small muscle that controls blood circulation in the capillary beds

stroke volume
: the volume of blood pumped into the aorta per contraction of the left ventricle

Introduction

class="introduction" Just as humans recycle what we can and dump the remains into landfills, our bodies use and recycle what they can and excrete the remaining waste products. Our bodies' complex systems have developed ways to treat waste and maintain a balanced internal environment. (credit: modification of work by Redwin Law)

The daily intake recommendation for human water consumption is eight to ten glasses of water. In order to achieve a healthy balance, the human body should excrete the eight to ten glasses of water every day. This occurs via the processes of urination, defecation, sweating and, to a small extent, respiration. The organs and tissues of the human body are soaked in fluids that are maintained at constant temperature, pH, and solute concentration, all crucial elements of homeostasis. The solutes in body fluids are mainly mineral salts and sugars, and osmotic regulation is the process by which the mineral salts and water are kept in

balance. Osmotic homeostasis is maintained despite the influence of external factors like temperature, diet, and weather conditions.

Osmoregulation and Osmotic Balance

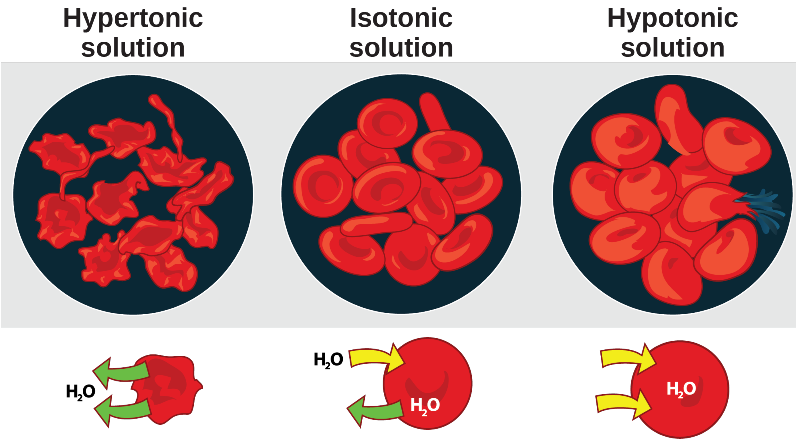
By the end of this section, you will be able to do the following:

- Define osmosis and explain its role within molecules
- Explain why osmoregulation and osmotic balance are important body functions
- Describe active transport mechanisms
- Explain osmolarity and the way in which it is measured
- Describe osmoregulators or osmoconformers and how these tools allow animals to adapt to different environments

Osmosis is the diffusion of water across a membrane in response to **osmotic pressure** caused by an imbalance of molecules on either side of the membrane. **Osmoregulation** is the process of maintenance of salt and water balance (**osmotic balance**) across membranes within the body's fluids, which are composed of water, plus electrolytes and non-electrolytes. An **electrolyte** is a solute that dissociates into ions when dissolved in water. A **non-electrolyte**, in contrast, doesn't dissociate into ions during water dissolution. Both electrolytes and non-electrolytes contribute to the osmotic balance. The body's fluids include blood plasma, the cytosol within cells, and interstitial fluid, the fluid that exists in the spaces between cells and tissues of the body. The membranes of the body

(such as the pleural, serous, and cell membranes) are **semi-permeable membranes**. Semi-permeable membranes are permeable (or permissive) to certain types of solutes and water. Solutions on two sides of a semi-permeable membrane tend to equalize in solute concentration by movement of solutes and/or water across the membrane. As seen in [link], a cell placed in water tends to swell due to gain of water from the hypotonic or “low salt” environment. A cell placed in a solution with higher salt concentration, on the other hand, tends to make the membrane shrivel up due to loss of water into the hypertonic or “high salt” environment. Isotonic cells have an equal concentration of solutes inside and outside the cell; this equalizes the osmotic pressure on either side of the cell membrane which is a semi-permeable membrane.

Cells placed in a hypertonic environment tend to shrink due to loss of water. In a hypotonic environment, cells tend to swell due to intake of water. The blood maintains an isotonic environment so that cells neither shrink nor swell. (credit: Mariana Ruiz Villareal)



The body does not exist in isolation. There is a constant input of water and electrolytes into the system. While osmoregulation is achieved across membranes within the body, excess electrolytes and wastes are transported to the kidneys and excreted, helping to maintain osmotic balance.

Need for Osmoregulation

Biological systems constantly interact and exchange water and nutrients with the environment by way of consumption of food and water and through excretion in the form of sweat, urine, and feces. Without a mechanism to regulate osmotic pressure, or when a disease damages this mechanism, there is a tendency to accumulate toxic waste and water, which can have dire consequences.

Mammalian systems have evolved to regulate not

only the overall osmotic pressure across membranes, but also specific concentrations of important electrolytes in the three major fluid compartments: blood plasma, extracellular fluid, and intracellular fluid. Since osmotic pressure is regulated by the movement of water across membranes, the volume of the fluid compartments can also change temporarily. Because blood plasma is one of the fluid components, osmotic pressures have a direct bearing on blood pressure.

Transport of Electrolytes across Cell Membranes

Electrolytes, such as sodium chloride, ionize in water, meaning that they dissociate into their component ions. In water, sodium chloride (NaCl), dissociates into the sodium ion (Na^+) and the chloride ion ($Cl^–$). The most important ions, whose concentrations are very closely regulated in body fluids, are the cations sodium (Na^+), potassium (K^+), calcium (Ca^{+2}), magnesium (Mg^{+2}), and the anions chloride (Cl^-), carbonate (CO_3^{-2}), bicarbonate (HCO_3^-), and phosphate(PO_3^-). Electrolytes are lost from the body during urination and perspiration. For this reason, athletes are encouraged to replace electrolytes and fluids during periods of increased activity and perspiration.

Osmotic pressure is influenced by the concentration

of solutes in a solution. It is directly proportional to the number of solute atoms or molecules and not dependent on the size of the solute molecules. Because electrolytes dissociate into their component ions, they, in essence, add more solute particles into the solution and have a greater effect on osmotic pressure, per mass than compounds that do not dissociate in water, such as glucose.

Water can pass through membranes by passive diffusion. If electrolyte ions could passively diffuse across membranes, it would be impossible to maintain specific concentrations of ions in each fluid compartment therefore they require special mechanisms to cross the semi-permeable membranes in the body. This movement can be accomplished by facilitated diffusion and active transport. Facilitated diffusion requires protein-based channels for moving the solute. Active transport requires energy in the form of ATP conversion, carrier proteins, or pumps in order to move ions against the concentration gradient.

Concept of Osmolality and Milliequivalent

In order to calculate osmotic pressure, it is necessary to understand how solute concentrations are measured. The unit for measuring solutes is the **mole**. One mole is defined as the gram molecular

weight of the solute. For example, the molecular weight of sodium chloride is 58.44. Thus, one mole of sodium chloride weighs 58.44 grams. The **molarity** of a solution is the number of moles of solute per liter of solution. The **molality** of a solution is the number of moles of solute per kilogram of solvent. If the solvent is water, one kilogram of water is equal to one liter of water. While molarity and molality are used to express the concentration of solutions, electrolyte concentrations are usually expressed in terms of milliequivalents per liter (mEq/L): the mEq/L is equal to the ion concentration (in millimoles) multiplied by the number of electrical charges on the ion. The unit of milliequivalent takes into consideration the ions present in the solution (since electrolytes form ions in aqueous solutions) and the charge on the ions.

Thus, for ions that have a charge of one, one milliequivalent is equal to one millimole. For ions that have a charge of two (like calcium), one milliequivalent is equal to 0.5 millimoles. Another unit for the expression of electrolyte concentration is the milliosmole (mOsm), which is the number of milliequivalents of solute per kilogram of solvent. Body fluids are usually maintained within the range of 280 to 300 mOsm.

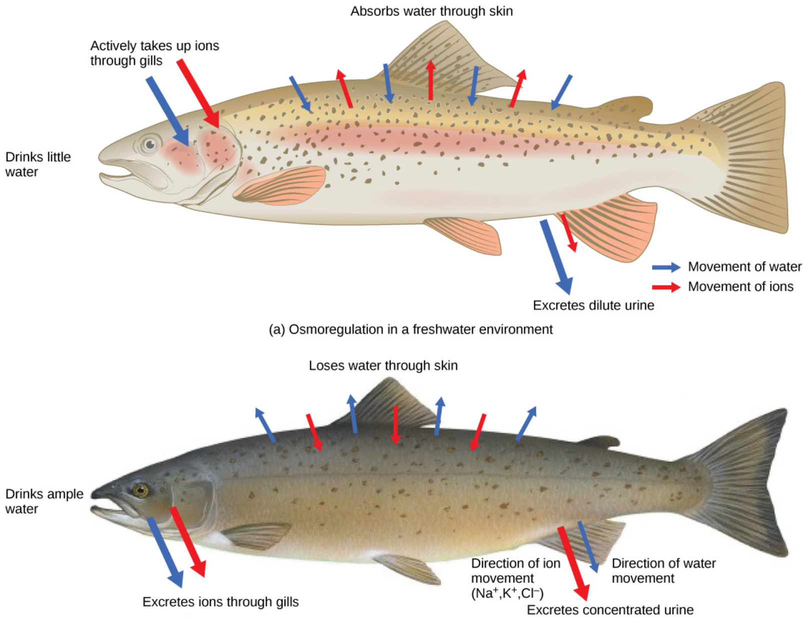
Fish are osmoregulators, but must use different mechanisms to survive in (a) freshwater or (b) saltwater environments. (credit: modification of

work by Duane Raver, NOAA)

Osmoregulators and Osmoconformers

Persons lost at sea without any freshwater to drink are at risk of severe dehydration because the human body cannot adapt to drinking seawater, which is hypertonic in comparison to body fluids. Organisms such as goldfish that can tolerate only a relatively narrow range of salinity are referred to as stenohaline. About 90 percent of all bony fish are restricted to either freshwater or seawater. They are incapable of osmotic regulation in the opposite environment. It is possible, however, for a few fishes like salmon to spend part of their life in freshwater and part in seawater. Organisms like the salmon and molly that can tolerate a relatively wide range of salinity are referred to as euryhaline organisms. This is possible because some fish have evolved **osmoregulatory** mechanisms to survive in all kinds of aquatic environments. When they live in freshwater, their bodies tend to take up water because the environment is relatively hypotonic, as illustrated in [link]**a**. In such hypotonic environments, these fish do not drink much water. Instead, they pass a lot of very dilute urine, and they achieve electrolyte balance by active transport of salts through the gills. When they move to a hypertonic marine environment, these fish start drinking seawater; they excrete the excess salts through their gills and their urine, as illustrated in

[link]**b**. Most marine invertebrates, on the other hand, may be isotonic with seawater (**osmoconformers**). Their body fluid concentrations conform to changes in seawater concentration. Cartilaginous fishes' salt composition of the blood is similar to bony fishes; however, the blood of sharks contains the organic compounds urea and trimethylamine oxide (TMAO). This does not mean that their electrolyte composition is similar to that of seawater. They achieve isotonicity with the sea by storing large concentrations of urea. These animals that secrete urea are called ureotelic animals. TMAO stabilizes proteins in the presence of high urea levels, preventing the disruption of peptide bonds that would occur in other animals exposed to similar levels of urea. Sharks are cartilaginous fish with a rectal gland to secrete salt and assist in osmoregulation.



(a) Osmoregulation in a freshwater environment

(b) Osmoregulation in a saltwater environment

Career Connection

Dialysis Technician

Dialysis is a medical process of removing wastes and excess water from the blood by diffusion and ultrafiltration. When kidney function fails, dialysis must be done to artificially rid the body of wastes. This is a vital process to keep patients alive. In some cases, the patients undergo artificial dialysis until they are eligible for a kidney transplant. In others who are not candidates for kidney transplants, dialysis is a life-long necessity. Dialysis technicians typically work in hospitals and clinics. While some roles in this field include equipment development and maintenance, most

dialysis technicians work in direct patient care. Their on-the-job duties, which typically occur under the direct supervision of a registered nurse, focus on providing dialysis treatments. This can include reviewing patient history and current condition, assessing and responding to patient needs before and during treatment, and monitoring the dialysis process. Treatment may include taking and reporting a patient's vital signs and preparing solutions and equipment to ensure accurate and sterile procedures.

Section Summary

Solute concentrations across semi-permeable membranes influence the movement of water and solutes across the membrane. It is the number of solute molecules and not the molecular size that is important in osmosis. Osmoregulation and osmotic balance are important bodily functions, resulting in water and salt balance. Not all solutes can pass through a semi-permeable membrane. Osmosis is the movement of water across the membrane. Osmosis occurs to equalize the number of solute molecules across a semi-permeable membrane by the movement of water to the side of higher solute concentration. Facilitated diffusion utilizes protein

channels to move solute molecules from areas of higher to lower concentration while active transport mechanisms are required to move solutes against concentration gradients. Osmolarity is measured in units of milliequivalents or milliosmoles, both of which take into consideration the number of solute particles and the charge on them. Fish that live in freshwater or saltwater adapt by being osmoregulators or osmoconformers.

Review Questions

When a dehydrated human patient needs to be given fluids intravenously, he or she is given:

1. water, which is hypotonic with respect to body fluids
2. saline at a concentration that is isotonic with respect to body fluids
3. glucose because it is a non-electrolyte
4. blood

B

The sodium ion is at the highest concentration in:

1. intracellular fluid
2. extracellular fluid
3. blood plasma
4. none of the above

B

Cells in a hypertonic solution tend to:

1. shrink due to water loss
2. swell due to water gain
3. stay the same size due to water moving into and out of the cell at the same rate
4. none of the above

A

Critical Thinking Questions

Why is excretion important in order to achieve osmotic balance?

Excretion allows an organism to rid itself of waste molecules that could be toxic if allowed to accumulate. It also allows the organism to

keep the amount of water and dissolved solutes in balance.

Why do electrolyte ions move across membranes by active transport?

Electrolyte ions often require special mechanisms to cross the semi-permeable membranes in the body. Active transport is the movement against a concentration gradient.

Glossary

electrolyte
: solute that breaks down into ions when dissolved in water

molality
: number of moles of solute per kilogram of solvent

molarity
: number of moles of solute per liter of solution

mole
: gram equivalent of the molecular weight of a substance

non-electrolyte

solute that does not break down into ions when dissolved in water

osmoconformer
: organism that changes its tonicity based on its environment

osmoregulation
: mechanism by which water and solute concentrations are maintained at desired levels

osmoregulator
: organism that maintains its tonicity irrespective of its environment

osmotic balance
: balance of the amount of water and salt input and output to and from a biological system without disturbing the desired osmotic pressure and solute concentration in every compartment

osmotic pressure
: pressure exerted on a membrane to equalize solute concentration on either side

semi-permeable membrane
: membrane that allows only certain solutes to pass through

The Kidneys and Osmoregulatory Organs
By the end of this section, you will be able to do the following:

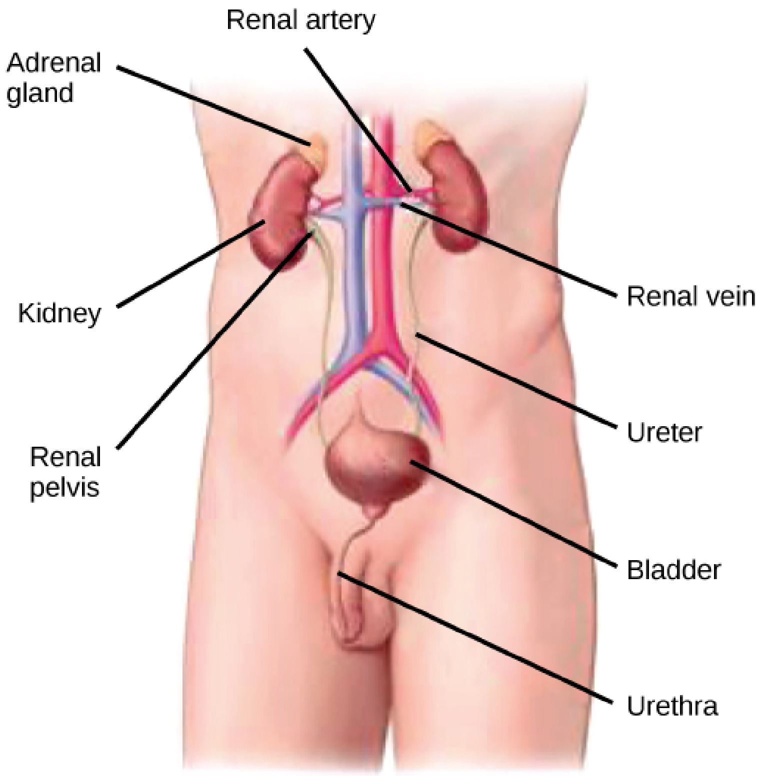
- Explain how the kidneys serve as the main osmoregulatory organs in mammalian systems
- Describe the structure of the kidneys and the functions of the parts of the kidney
- Describe how the nephron is the functional unit of the kidney and explain how it actively filters blood and generates urine
- Detail the three steps in the formation of urine: glomerular filtration, tubular reabsorption, and tubular secretion

Although the kidneys are the major osmoregulatory organ, the skin and lungs also play a role in the process. Water and electrolytes are lost through sweat glands in the skin, which helps moisturize and cool the skin surface, while the lungs expel a small amount of water in the form of mucous secretions and via evaporation of water vapor.
Kidneys filter the blood, producing urine that is stored in the bladder prior to elimination through the urethra. (credit: modification of work by NCI)

Kidneys: The Main Osmoregulatory Organ

The **kidneys**, illustrated in [link], are a pair of

bean-shaped structures that are located just below and posterior to the liver in the peritoneal cavity. The adrenal glands sit on top of each kidney and are also called the suprarenal glands. Kidneys filter blood and purify it. All the blood in the human body is filtered many times a day by the kidneys; these organs use up almost 25 percent of the oxygen absorbed through the lungs to perform this function. Oxygen allows the kidney cells to efficiently manufacture chemical energy in the form of ATP through aerobic respiration. The filtrate coming out of the kidneys is called **urine**.

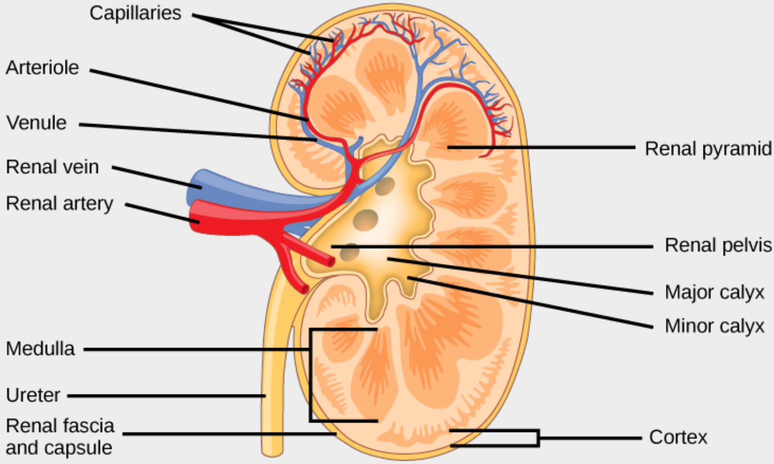


Kidney Structure

Externally, the kidneys are surrounded by three layers, illustrated in [link]. The outermost layer is a tough connective tissue layer called the **renal fascia**. The second layer is called the **perirenal fat capsule**, which helps anchor the kidneys in place. The third and innermost layer is the **renal capsule**. Internally, the kidney has three regions—an outer **cortex**, a **medulla** in the middle, and the **renal pelvis** in the region called the **hilum** of the kidney. The hilum is the concave part of the bean-shape where blood vessels and nerves enter and exit the kidney; it is also the point of exit for the ureters. The renal cortex is granular due to the presence of **nephrons**—the functional unit of the kidney. The medulla consists of multiple pyramidal tissue masses, called the **renal pyramids**. In between the pyramids are spaces called **renal columns** through which the blood vessels pass. The tips of the pyramids, called renal papillae, point toward the renal pelvis. There are, on average, eight renal pyramids in each kidney. The renal pyramids along with the adjoining cortical region are called the **lobes of the kidney**. The renal pelvis leads to the **ureter** on the outside of the kidney. On the inside of the kidney, the renal pelvis branches out into two or three extensions called the major **calyces**, which further branch into the minor calyces. The ureters are urine-bearing tubes that exit the kidney and empty into the **urinary bladder**.

Visual Connection

The internal structure of the kidney is shown. (credit: modification of work by NCI)



Which of the following statements about the kidney is false?

1. The renal pelvis drains into the ureter.
2. The renal pyramids are in the medulla.
3. The cortex covers the capsule.
4. Nephrons are in the renal cortex.

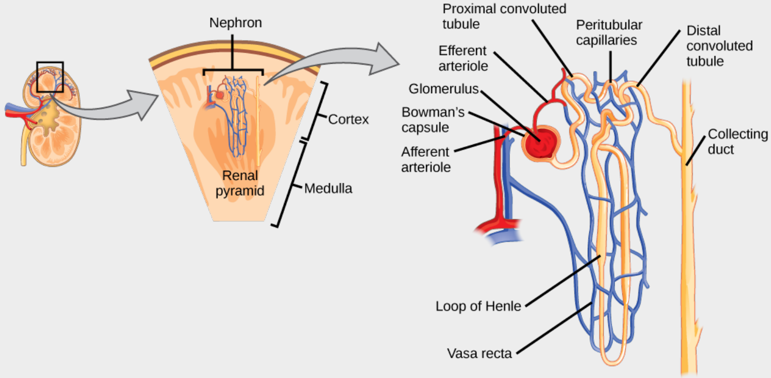
Because the kidney filters blood, its network of blood vessels is an important component of its structure and function. The arteries, veins, and nerves that supply the kidney enter and exit at the renal hilum. Renal blood supply starts with the branching of the aorta into the **renal arteries** (which are each named based on the region of the kidney they pass through) and ends with the exiting

of the **renal veins** to join the **inferior vena cava**. The renal arteries split into several **segmental arteries** upon entering the kidneys. Each segmental artery splits further into several **interlobar arteries** and enters the renal columns, which supply the renal lobes. The interlobar arteries split at the junction of the renal cortex and medulla to form the **arcuate arteries**. The arcuate "bow shaped" arteries form arcs along the base of the medullary pyramids. **Cortical radiate arteries**, as the name suggests, radiate out from the arcuate arteries. The cortical radiate arteries branch into numerous afferent arterioles, and then enter the capillaries supplying the nephrons. Veins trace the path of the arteries and have similar names, except there are no segmental veins.

As mentioned previously, the functional unit of the kidney is the nephron, illustrated in [link]. Each kidney is made up of over one million nephrons that dot the renal cortex, giving it a granular appearance when sectioned sagittally. There are two types of nephrons—**cortical nephrons** (85 percent), which are deep in the renal cortex, and **juxtamedullary nephrons** (15 percent), which lie in the renal cortex close to the renal medulla. A nephron consists of three parts—a **renal corpuscle**, a **renal tubule**, and the associated capillary network, which originates from the cortical radiate arteries.

Visual Connection

The nephron is the functional unit of the kidney. The glomerulus and convoluted tubules are located in the kidney cortex, while collecting ducts are located in the pyramids of the medulla. (credit: modification of work by NIDDK)



Which of the following statements about the nephron is false?

1. The collecting duct empties into the distal convoluted tubule.
2. The Bowman's capsule surrounds the glomerulus.
3. The loop of Henle is between the proximal and distal convoluted tubules.
4. The loop of Henle empties into the distal convoluted tubule.

Renal Corpuscle

The renal corpuscle, located in the renal cortex, is

made up of a network of capillaries known as the **glomerulus** and the capsule, a cup-shaped chamber that surrounds it, called the glomerular or **Bowman's capsule**.

Renal Tubule

The renal tubule is a long and convoluted structure that emerges from the glomerulus and can be divided into three parts based on function. The first part is called the **proximal convoluted tubule (PCT)** due to its proximity to the glomerulus; it stays in the renal cortex. The second part is called the **loop of Henle**, or nephritic loop, because it forms a loop (with **descending** and **ascending limbs**) that goes through the renal medulla. The third part of the renal tubule is called the **distal convoluted tubule (DCT)** and this part is also restricted to the renal cortex. The DCT, which is the last part of the nephron, connects and empties its contents into collecting ducts that line the medullary pyramids. The collecting ducts amass contents from multiple nephrons and fuse together as they enter the papillae of the renal medulla.

Capillary Network within the Nephron

The capillary network that originates from the renal arteries supplies the nephron with blood that needs to be filtered. The branch that enters the glomerulus is called the **afferent arteriole**. The branch that

exits the glomerulus is called the **efferent arteriole**. Within the glomerulus, the network of capillaries is called the glomerular capillary bed. Once the efferent arteriole exits the glomerulus, it forms the **peritubular capillary network**, which surrounds and interacts with parts of the renal tubule. In cortical nephrons, the peritubular capillary network surrounds the PCT and DCT. In juxtamedullary nephrons, the peritubular capillary network forms a network around the loop of Henle and is called the **vasa recta**.

Link to Learning

Go to this website to see another coronal section of the kidney and to explore an animation of the workings of nephrons.

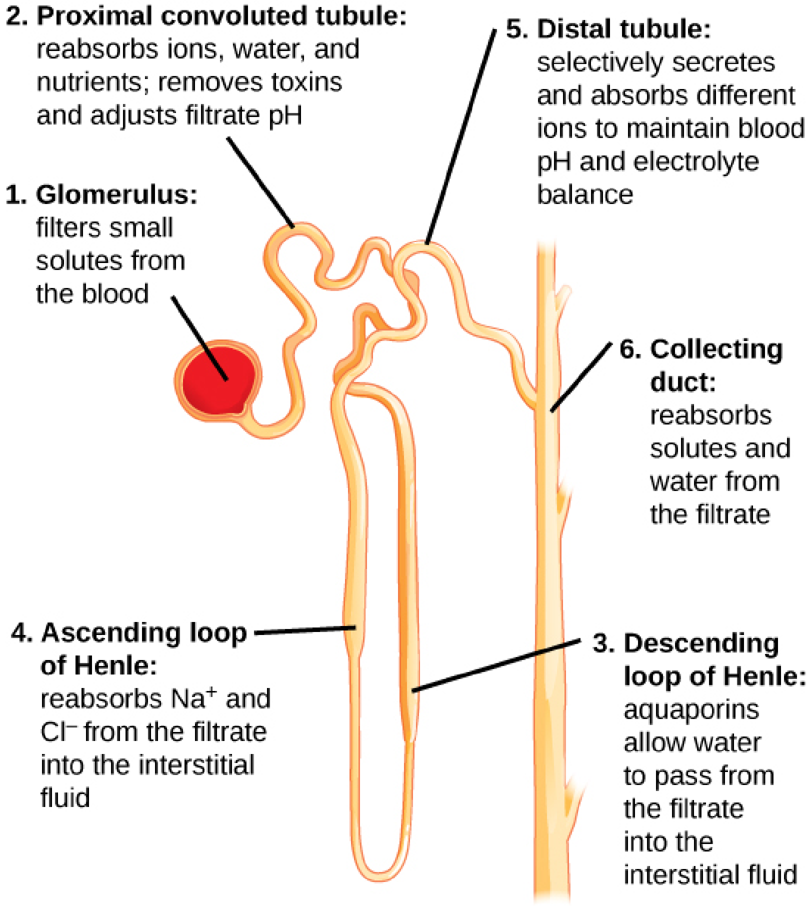
Each part of the nephron performs a different function in filtering waste and maintaining homeostatic balance. (1) The glomerulus forces small solutes out of the blood by pressure. (2) The proximal convoluted tubule reabsorbs ions, water, and nutrients from the filtrate into the interstitial fluid, and actively transports toxins and drugs from the interstitial fluid into the filtrate. The proximal convoluted tubule also adjusts blood pH by selectively secreting ammonia (NH_3) into the

filtrate, where it reacts with H^+ to form NH_4^+. The more acidic the filtrate, the more ammonia is secreted. (3) The descending loop of Henle is lined with cells containing aquaporins that allow water to pass from the filtrate into the interstitial fluid. (4) In the thin part of the ascending loop of Henle, Na^+ and Cl^- ions diffuse into the interstitial fluid. In the thick part, these same ions are actively transported into the interstitial fluid. Because salt but not water is lost, the filtrate becomes more dilute as it travels up the limb. (5) In the distal convoluted tubule, K^+ and H^+ ions are selectively secreted into the filtrate, while Na^+, Cl^-, and HCO_3^- ions are reabsorbed to maintain pH and electrolyte balance in the blood. (6) The collecting duct reabsorbs solutes and water from the filtrate, forming dilute urine. (credit: modification of work by NIDDK)

Kidney Function and Physiology

Kidneys filter blood in a three-step process. First, the nephrons filter blood that runs through the capillary network in the glomerulus. Almost all solutes, except for proteins, are filtered out into the glomerulus by a process called **glomerular filtration**. Second, the filtrate is collected in the renal tubules. Most of the solutes get reabsorbed in the PCT by a process called **tubular reabsorption**. In the loop of Henle, the filtrate continues to exchange solutes and water with the renal medulla and the peritubular capillary network. Water is also

reabsorbed during this step. Then, additional solutes and wastes are secreted into the kidney tubules during **tubular secretion**, which is, in essence, the opposite process to tubular reabsorption. The collecting ducts collect filtrate coming from the nephrons and fuse in the medullary papillae. From here, the papillae deliver the filtrate, now called urine, into the minor calyces that eventually connect to the ureters through the renal pelvis. This entire process is illustrated in [link].



Glomerular Filtration

Glomerular filtration filters out most of the solutes due to high blood pressure and specialized membranes in the afferent arteriole. The blood pressure in the glomerulus is maintained independent of factors that affect systemic blood pressure. The "leaky" connections between the endothelial cells of the glomerular capillary network allow solutes to pass through easily. All solutes in the glomerular capillaries, except for macromolecules like proteins, pass through by passive diffusion. There is no energy requirement at this stage of the filtration process. **Glomerular filtration rate (GFR)** is the volume of glomerular filtrate formed per minute by the kidneys. GFR is regulated by multiple mechanisms and is an important indicator of kidney function.

> Link to Learning
> To learn more about the vascular system of kidneys, click through this review and the steps of blood flow.

Tubular Reabsorption and Secretion

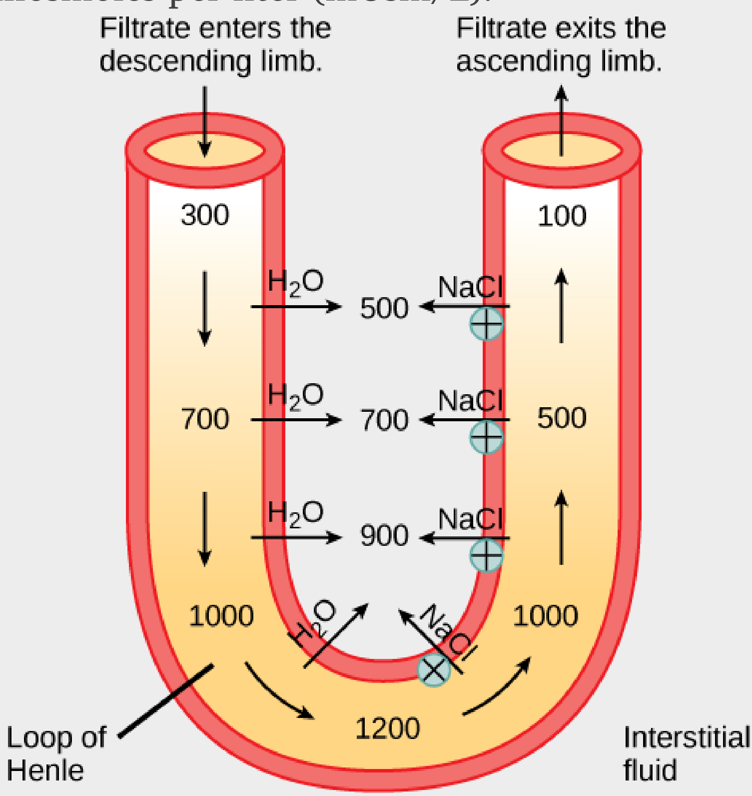
Tubular reabsorption occurs in the PCT part of the

renal tubule. Almost all nutrients are reabsorbed, and this occurs either by passive or active transport. Reabsorption of water and some key electrolytes are regulated and can be influenced by hormones. Sodium (Na^+) is the most abundant ion and most of it is reabsorbed by active transport and then transported to the peritubular capillaries. Because Na^+ is actively transported out of the tubule, water follows it to even out the osmotic pressure. Water is also independently reabsorbed into the peritubular capillaries due to the presence of aquaporins, or water channels, in the PCT. This occurs due to the low blood pressure and high osmotic pressure in the peritubular capillaries. However, every solute has a **transport maximum** and the excess is not reabsorbed.

In the loop of Henle, the permeability of the membrane changes. The descending limb is permeable to water, not solutes; the opposite is true for the ascending limb. Additionally, the loop of Henle invades the renal medulla, which is naturally high in salt concentration and tends to absorb water from the renal tubule and concentrate the filtrate. The osmotic gradient increases as it moves deeper into the medulla. Because two sides of the loop of Henle perform opposing functions, as illustrated in [link], it acts as a **countercurrent multiplier**. The vasa recta around it acts as the **countercurrent exchanger**.

Visual Connection

The loop of Henle acts as a countercurrent multiplier that uses energy to create concentration gradients. The descending limb is water permeable. Water flows from the filtrate to the interstitial fluid, so osmolality inside the limb increases as it descends into the renal medulla. At the bottom, the osmolality is higher inside the loop than in the interstitial fluid. Thus, as filtrate enters the ascending limb, Na+ and Cl- ions exit through ion channels present in the cell membrane. Further up, Na+ is actively transported out of the filtrate and Cl- follows. Osmolarity is given in units of milliosmoles per liter (mOsm/L).



Loop diuretics are drugs sometimes used to treat hypertension. These drugs inhibit the reabsorption of Na^+ and Cl^- ions by the ascending limb of the loop of Henle. A side effect is that they increase urination. Why do you think this is the case?

By the time the filtrate reaches the DCT, most of the urine and solutes have been reabsorbed. If the body requires additional water, all of it can be reabsorbed at this point. Further reabsorption is controlled by hormones, which will be discussed in a later section. Excretion of wastes occurs due to lack of reabsorption combined with tubular secretion. Undesirable products like metabolic wastes, urea, uric acid, and certain drugs, are excreted by tubular secretion. Most of the tubular secretion happens in the DCT, but some occurs in the early part of the collecting duct. Kidneys also maintain an acid-base balance by secreting excess H^+ ions.

Although parts of the renal tubules are named proximal and distal, in a cross-section of the kidney, the tubules are placed close together and in contact with each other and the glomerulus. This allows for exchange of chemical messengers between the different cell types. For example, the DCT ascending limb of the loop of Henle has masses of cells called **macula densa**, which are in contact with cells of the afferent arterioles called **juxtaglomerular cells**.

Together, the macula densa and juxtaglomerular cells form the juxtaglomerular complex (JGC). The JGC is an endocrine structure that secretes the enzyme renin and the hormone erythropoietin. When hormones trigger the macula densa cells in the DCT due to variations in blood volume, blood pressure, or electrolyte balance, these cells can immediately communicate the problem to the capillaries in the afferent and efferent arterioles, which can constrict or relax to change the glomerular filtration rate of the kidneys.

Career Connection

Nephrologist

A nephrologist studies and deals with diseases of the kidneys—both those that cause kidney failure (such as diabetes) and the conditions that are produced by kidney disease (such as hypertension). Blood pressure, blood volume, and changes in electrolyte balance come under the purview of a nephrologist.

Nephrologists usually work with other physicians who refer patients to them or consult with them about specific diagnoses and treatment plans. Patients are usually referred to a nephrologist for symptoms such as blood or protein in the urine, very high blood pressure, kidney stones, or renal failure.

Nephrology is a subspecialty of internal medicine.

To become a nephrologist, medical school is followed by additional training to become certified in internal medicine. An additional two or more years is spent specifically studying kidney disorders and their accompanying effects on the body.

Section Summary

The kidneys are the main osmoregulatory organs in mammalian systems; they function to filter blood and maintain the osmolarity of body fluids at 300 mOsm. They are surrounded by three layers and are made up internally of three distinct regions—the cortex, medulla, and pelvis.

The blood vessels that transport blood into and out of the kidneys arise from and merge with the aorta and inferior vena cava, respectively. The renal arteries branch out from the aorta and enter the kidney where they further divide into segmental, interlobar, arcuate, and cortical radiate arteries.

The nephron is the functional unit of the kidney, which actively filters blood and generates urine. The nephron is made up of the renal corpuscle and renal tubule. Cortical nephrons are found in the renal cortex, while juxtamedullary nephrons are found in

the renal cortex close to the renal medulla. The nephron filters and exchanges water and solutes with two sets of blood vessels and the tissue fluid in the kidneys.

There are three steps in the formation of urine: glomerular filtration, which occurs in the glomerulus; tubular reabsorption, which occurs in the renal tubules; and tubular secretion, which also occurs in the renal tubules.

Visual Connection Questions

[link] Which of the following statements about the kidney is false?

1. The renal pelvis drains into the ureter.
2. The renal pyramids are in the medulla.
3. The cortex covers the capsule.
4. Nephrons are in the renal cortex.

[link] C

[link] Which of the following statements about the nephron is false?

1. The collecting duct empties into the distal

convoluted tubule.
2. The Bowman's capsule surrounds the glomerulus.
3. The loop of Henle is between the proximal and distal convoluted tubules.
4. The loop of Henle empties into the distal convoluted tubule.

[link] A

[link] Loop diuretics are drugs sometimes used to treat hypertension. These drugs inhibit the reabsorption of Na^+ and Cl^- ions by the ascending limb of the loop of Henle. A side effect is that they increase urination. Why do you think this is the case?

[link] Loop diuretics decrease the excretion of salt into the renal medulla, thereby reducing its osmolality. As a result, less water is excreted into the medulla by the descending limb, and more water is excreted as urine.

Review Questions

The macula densa is/are:

1. present in the renal medulla.
2. dense tissue present in the outer layer of the kidney.
3. cells present in the DCT and collecting tubules.
4. present in blood capillaries.

C

The osmolarity of body fluids is maintained at ______.

1. 100 mOsm
2. 300 mOsm
3. 1000 mOsm
4. it is not constantly maintained

B

The gland located at the top of the kidney is the ______ gland.

1. adrenal
2. pituitary
3. thyroid
4. thymus

A

Critical Thinking Questions

Why are the loop of Henle and vasa recta important for the formation of concentrated urine?

The loop of Henle is part of the renal tubule that loops into the renal medulla. In the loop of Henle, the filtrate exchanges solutes and water with the renal medulla and the vasa recta (the peritubular capillary network). The vasa recta acts as the countercurrent exchanger. The kidneys maintain the osmolality of the rest of the body at a constant 300 mOsm by concentrating the filtrate as it passes through the loop of Henle.

Describe the structure of the kidney.

Externally, the kidneys are surrounded by three layers. The outermost layer is a tough connective tissue layer called the renal fascia. The second layer is called the perirenal fat

capsule, which helps anchor the kidneys in place. The third and innermost layer is the renal capsule. Internally, the kidney has three regions—an outer cortex, a medulla in the middle, and the renal pelvis in the region called the hilum of the kidney, which is the concave part of the “bean” shape.

Glossary

afferent arteriole
: arteriole that branches from the cortical radiate artery and enters the glomerulus

arcuate artery
: artery that branches from the interlobar artery and arches over the base of the renal pyramids

ascending limb
: part of the loop of Henle that ascends from the renal medulla to the renal cortex

Bowman's capsule
: structure that encloses the glomerulus

calyx
: structure that connects the renal pelvis to the renal medulla

cortex (animal)

outer layer of an organ like the kidney or adrenal gland

cortical nephron
: nephron that lies in the renal cortex

cortical radiate artery
: artery that radiates from the arcuate arteries into the renal cortex

countercurrent exchanger
: peritubular capillary network that allows exchange of solutes and water from the renal tubules

countercurrent multiplier
: osmotic gradient in the renal medulla that is responsible for concentration of urine

descending limb
: part of the loop of Henle that descends from the renal cortex into the renal medulla

distal convoluted tubule (DCT)
: part of the renal tubule that is the most distant from the glomerulus

efferent arteriole
: arteriole that exits from the glomerulus

glomerular filtration
: filtration of blood in the glomerular capillary network into the glomerulus

glomerular filtration rate (GFR)
: amount of filtrate formed by the glomerulus per minute

glomerulus (renal)
: part of the renal corpuscle that contains the capillary network

hilum
: region in the renal pelvis where blood vessels, nerves, and ureters bunch before entering or exiting the kidney

inferior vena cava
: one of the main veins in the human body

interlobar artery
: artery that branches from the segmental artery and travels in between the renal lobes

juxtaglomerular cell
: cell in the afferent and efferent arterioles that responds to stimuli from the macula densa

juxtamedullary nephron
: nephron that lies in the cortex but close to the renal medulla

kidney
: organ that performs excretory and osmoregulatory functions

lobes of the kidney

renal pyramid along with the adjoining cortical region

loop of Henle
: part of the renal tubule that loops into the renal medulla

macula densa
: group of cells that senses changes in sodium ion concentration; present in parts of the renal tubule and collecting ducts

medulla
: middle layer of an organ like the kidney or adrenal gland

nephron
: functional unit of the kidney

perirenal fat capsule
: fat layer that suspends the kidneys

peritubular capillary network
: capillary network that surrounds the renal tubule after the efferent artery exits the glomerulus

proximal convoluted tubule (PCT)
: part of the renal tubule that lies close to the glomerulus

renal artery
: branch of the artery that enters the kidney

renal capsule
: layer that encapsulates the kidneys

renal column
: area of the kidney through which the interlobar arteries travel in the process of supplying blood to the renal lobes

renal corpuscle
: glomerulus and the Bowman's capsule together

renal fascia
: connective tissue that supports the kidneys

renal pelvis
: region in the kidney where the calyces join the ureters

renal pyramid
: conical structure in the renal medulla

renal tubule
: tubule of the nephron that arises from the glomerulus

renal vein
: branch of a vein that exits the kidney and joins the inferior vena cava

segmental artery
: artery that branches from the renal artery

transport maximum
: maximum amount of solute that can be transported out of the renal tubules during reabsorption

tubular reabsorption
: reclamation of water and solutes that got filtered out in the glomerulus

tubular secretion
: process of secretion of wastes that do not get reabsorbed

ureter
: urine-bearing tube coming out of the kidney; carries urine to the bladder

urinary bladder
: structure that the ureters empty the urine into; stores urine

urine
: filtrate produced by kidneys that gets excreted out of the body

vasa recta
: peritubular network that surrounds the loop of Henle of the juxtamedullary nephrons

Excretion Systems

By the end of this section, you will be able to do the following:

- Explain how vacuoles, present in microorganisms, work to excrete waste
- Describe the way in which flame cells and nephridia in worms perform excretory functions and maintain osmotic balance
- Explain how insects use Malpighian tubules to excrete wastes and maintain osmotic balance

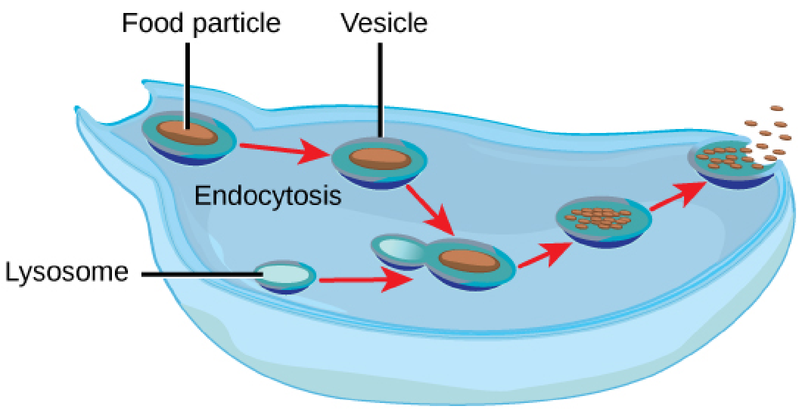
Microorganisms and invertebrate animals use more primitive and simple mechanisms to get rid of their metabolic wastes than the mammalian system of kidney and urinary function. Three excretory systems evolved in organisms before complex kidneys: vacuoles, flame cells, and Malpighian tubules.

Some unicellular organisms, such as the amoeba, ingest food by endocytosis. The food vesicle fuses with a lysosome, which digests the food. Waste is excreted by exocytosis.

Contractile Vacuoles in Microorganisms

The most fundamental feature of life is the presence of a cell. In other words, a cell is the simplest functional unit of a life. Bacteria are unicellular, prokaryotic organisms that have some of the least

complex life processes in place; however, prokaryotes such as bacteria do not contain membrane-bound vacuoles. The cells of microorganisms like bacteria, protozoa, and fungi are bound by cell membranes and use them to interact with the environment. Some cells, including some leucocytes in humans, are able to engulf food by endocytosis—the formation of vesicles by involution of the cell membrane within the cells. The same vesicles are able to interact and exchange metabolites with the intracellular environment. In some unicellular eukaryotic organisms such as the amoeba, shown in [link], cellular wastes and excess water are excreted by exocytosis, when the contractile vacuoles merge with the cell membrane and expel wastes into the environment. Contractile vacuoles (CV) should not be confused with vacuoles, which store food or water.



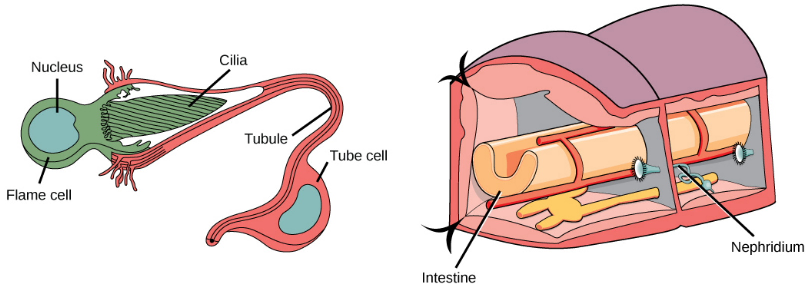
In the excretory system of the (a) planaria, cilia of flame cells propel waste through a tubule formed by a tube cell. Tubules are connected into branched structures that lead to pores located all along the

sides of the body. The filtrate is secreted through these pores. In (b) annelids such as earthworms, nephridia filter fluid from the coelom, or body cavity. Beating cilia at the opening of the nephridium draw water from the coelom into a tubule. As the filtrate passes down the tubules, nutrients and other solutes are reabsorbed by capillaries. Filtered fluid containing nitrogenous and other wastes is stored in a bladder and then secreted through a pore in the side of the body.

Flame Cells of Planaria and Nephridia of Worms

As multicellular systems evolved to have organ systems that divided the metabolic needs of the body, individual organs evolved to perform the excretory function. Planaria are flatworms that live in freshwater. Their excretory system consists of two tubules connected to a highly branched duct system. The cells in the tubules are called **flame cells** (or **protonephridia**) because they have a cluster of cilia that looks like a flickering flame when viewed under the microscope, as illustrated in [link]**a**. The cilia propel waste matter down the tubules and out of the body through excretory pores that open on the body surface; cilia also draw water from the interstitial fluid, allowing for filtration. Any valuable metabolites are recovered by reabsorption. Flame cells are found in flatworms, including parasitic tapeworms and free-living planaria. They also

maintain the organism’s osmotic balance.



(a) Flame cell of a planarian

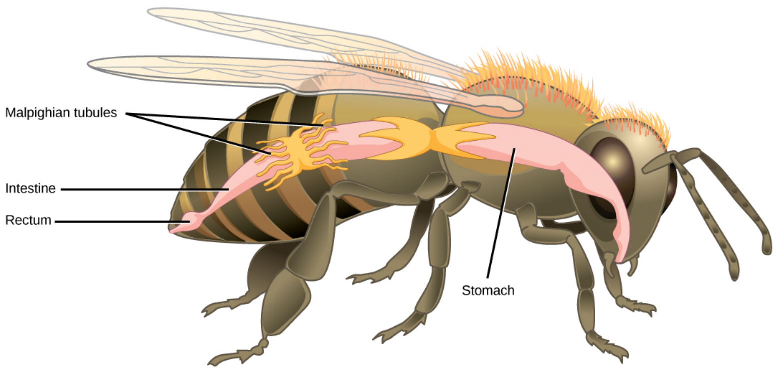
(b) Nephridium of an earthworm

Earthworms (annelids) have slightly more evolved excretory structures called **nephridia**, illustrated in [link]**b**. A pair of nephridia is present on each segment of the earthworm. They are similar to flame cells in that they have a tubule with cilia. Excretion occurs through a pore called the **nephridiopore**. They are more evolved than the flame cells in that they have a system for tubular reabsorption by a capillary network before excretion.

Malpighian tubules of insects and other terrestrial arthropods remove nitrogenous wastes and other solutes from the hemolymph. Na^+ and/or K^+ ions are actively transported into the lumen of the tubules. Water then enters the tubules via osmosis, forming urine. The urine passes through the intestine, and into the rectum. There, nutrients diffuse back into the hemolymph. Na^+ and/or K^+ ions are pumped into the hemolymph, and water follows. The concentrated waste is then excreted.

Malpighian Tubules of Insects

Malpighian tubules are found lining the gut of some species of arthropods, such as the bee illustrated in [link]. They are usually found in pairs and the number of tubules varies with the species of insect. Malpighian tubules are convoluted, which increases their surface area, and they are lined with **microvilli** for reabsorption and maintenance of osmotic balance. Malpighian tubules work cooperatively with specialized glands in the wall of the rectum. Body fluids are not filtered as in the case of nephridia; urine is produced by tubular secretion mechanisms by the cells lining the Malpighian tubules that are bathed in hemolymph (a mixture of blood and interstitial fluid that is found in insects and other arthropods as well as most mollusks). Metabolic wastes like uric acid freely diffuse into the tubules. There are exchange pumps lining the tubules, which actively transport H^+ ions into the cell and K^+ or Na^+ ions out; water passively follows to form urine. The secretion of ions alters the osmotic pressure which draws water, electrolytes, and nitrogenous waste (uric acid) into the tubules. Water and electrolytes are reabsorbed when these organisms are faced with low-water environments, and uric acid is excreted as a thick paste or powder. Not dissolving wastes in water helps these organisms to conserve water; this is especially important for life in dry environments.



Link to Learning

See a dissected cockroach, including a close-up look at its Malpighian tubules, in this video.

Section Summary

Many systems have evolved for excreting wastes that are simpler than the kidney and urinary systems of vertebrate animals. The simplest system is that of contractile vacuoles present in microorganisms. Flame cells and nephridia in worms perform excretory functions and maintain osmotic balance. Some insects have evolved Malpighian tubules to excrete wastes and maintain osmotic balance.

Review Questions

Active transport of K+ in Malpighian tubules ensures that:

1. water follows K+ to make urine
2. osmotic balance is maintained between waste matter and bodily fluids
3. both a and b
4. neither a nor b

C

Contractile vacuoles in microorganisms:

1. exclusively perform an excretory function
2. can perform many functions, one of which is excretion of metabolic wastes
3. originate from the cell membrane
4. both b and c

D

Flame cells are primitive excretory organs

found in _____.

1. arthropods
2. annelids
3. mammals
4. flatworms

D

Critical Thinking Questions

Why might specialized organs have evolved for excretion of wastes?

The removal of wastes, which could otherwise be toxic to an organism, is extremely important for survival. Having organs that specialize in this process and that operate separately from other organs provides a measure of safety for the organism.

Explain two different excretory systems other than the kidneys.

---

(1) Microorganisms engulf food by endocytosis

—the formation of vacuoles by involution of the cell membrane within the cells. The same vacuoles interact and exchange metabolites with the intracellular environment. Cellular wastes are excreted by exocytosis when the vacuoles merge with the cell membrane and excrete wastes into the environment. (2) Flatworms have an excretory system that consists of two tubules. The cells in the tubules are called flame cells; they have a cluster of cilia that propel waste matter down the tubules and out of the body. (3) Annelids have nephridia which have a tubule with cilia. Excretion occurs through a pore called the nephridiopore. Annelids have a system for tubular reabsorption by a capillary network before excretion. (4) Malpighian tubules are found in some species of arthropods. They are usually found in pairs, and the number of tubules varies with the species of insect. Malpighian tubules are convoluted, which increases their surface area, and they are lined with microvilli for reabsorption and maintenance of osmotic balance. Metabolic wastes like uric acid freely diffuse into the tubules. Potassium ion pumps line the tubules, which actively transport out $K+$ ions, and water follows to form urine. Water and electrolytes are reabsorbed when these organisms are faced with low-water environments, and uric acid is excreted as a thick paste or powder. By not

dissolving wastes in water, these organisms conserve water.

Glossary

flame cell
: (also, protonephridia) excretory cell found in flatworms

Malpighian tubule
: excretory tubules found in arthropods

microvilli
: cellular processes that increase the surface area of cells

nephridia
: excretory structures found in annelids

nephridiopore
: pore found at the end of nephridia

Nitrogenous Wastes

By the end of this section, you will be able to do the following:

- Compare and contrast the way in which aquatic animals and terrestrial animals can eliminate toxic ammonia from their systems
- Compare the major byproduct of ammonia metabolism in vertebrate animals to that of birds, insects, and reptiles

Of the four major macromolecules in biological systems, both proteins and nucleic acids contain nitrogen. During the catabolism, or breakdown, of nitrogen-containing macromolecules, carbon, hydrogen, and oxygen are extracted and stored in the form of carbohydrates and fats. Excess nitrogen is excreted from the body. Nitrogenous wastes tend to form toxic **ammonia**, which raises the pH of body fluids. The formation of ammonia itself requires energy in the form of ATP and large quantities of water to dilute it out of a biological system. Animals that live in aquatic environments tend to release ammonia into the water. Animals that excrete ammonia are said to be **ammonotelic**. Terrestrial organisms have evolved other mechanisms to excrete nitrogenous wastes. The animals must detoxify ammonia by converting it into a relatively nontoxic form such as urea or uric acid. Mammals, including humans, produce urea, whereas reptiles and many terrestrial invertebrates

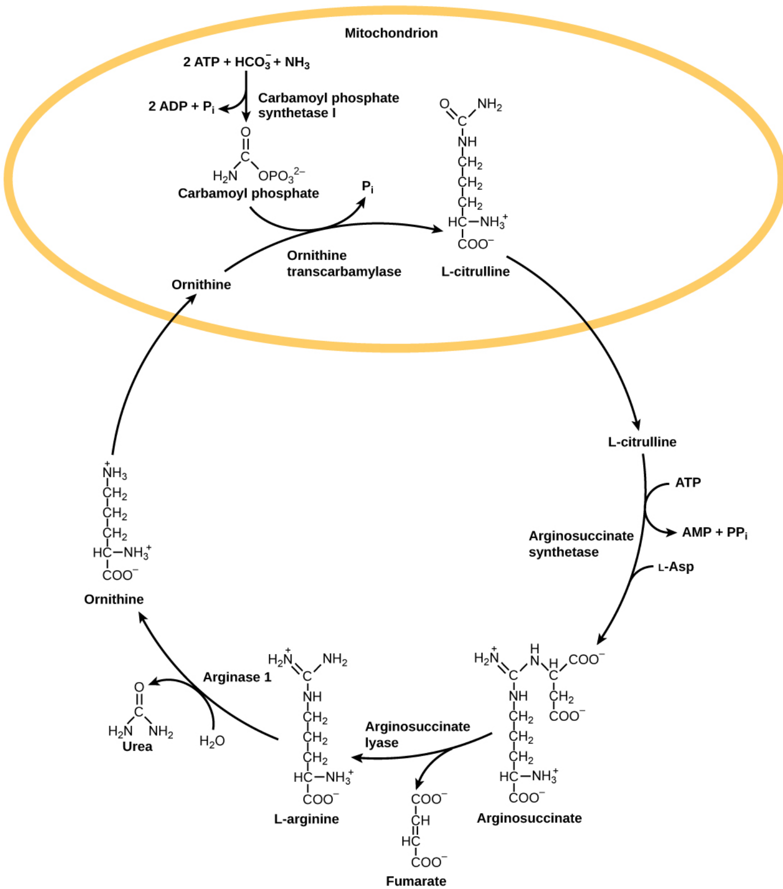
produce uric acid. Animals that secrete urea as the primary nitrogenous waste material are called **ureotelic** animals.

The urea cycle converts ammonia to urea.

Nitrogenous Waste in Terrestrial Animals: The Urea Cycle

The **urea cycle** is the primary mechanism by which mammals convert ammonia to urea. Urea is made in the liver and excreted in urine. The overall chemical reaction by which ammonia is converted to urea is $2\ NH_3$ (ammonia) $+\ CO_2 + 3\ ATP + H_2O \rightarrow H_2N\text{-}CO\text{-}NH_2$ (urea) $+\ 2\ ADP + 4\ P_i + AMP$.

The urea cycle utilizes five intermediate steps, catalyzed by five different enzymes, to convert ammonia to urea, as shown in [link]. The amino acid L-ornithine gets converted into different intermediates before being regenerated at the end of the urea cycle. Hence, the urea cycle is also referred to as the ornithine cycle. The enzyme ornithine transcarbamylase catalyzes a key step in the urea cycle and its deficiency can lead to accumulation of toxic levels of ammonia in the body. The first two reactions occur in the mitochondria and the last three reactions occur in the cytosol. Urea concentration in the blood, called **blood urea nitrogen** or BUN, is used as an indicator of kidney function.



Evolution Connection

Excretion of Nitrogenous Waste

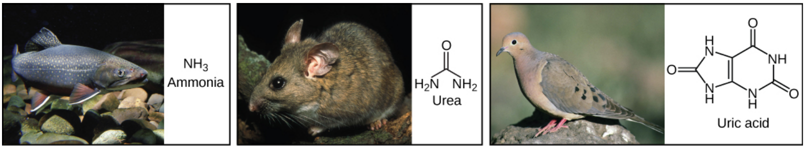
The theory of evolution proposes that life started in an aquatic environment. It is not surprising to see that biochemical pathways like the urea cycle evolved to adapt to a changing environment when terrestrial life forms evolved. Arid conditions probably led to the evolution of the uric acid

pathway as a means of conserving water.

Nitrogenous waste is excreted in different forms by different species. These include (a) ammonia, (b) urea, and (c) uric acid. (credit a: modification of work by Eric Engbretson, USFWS; credit b: modification of work by B. "Moose" Peterson, USFWS; credit c: modification of work by Dave Menke, USFWS)

Nitrogenous Waste in Birds and Reptiles: Uric Acid

Birds, reptiles, and most terrestrial arthropods convert toxic ammonia to **uric acid** or the closely related compound guanine (guano) instead of urea. Mammals also form some uric acid during breakdown of nucleic acids. Uric acid is a compound similar to purines found in nucleic acids. It is water insoluble and tends to form a white paste or powder; it is excreted by birds, insects, and reptiles. Conversion of ammonia to uric acid requires more energy and is much more complex than conversion of ammonia to urea [link].



(a) Many invertebrates and aquatic species excrete ammonia.

(b) Mammals, many adult amphibians, and some marine species excrete urea.

(c) Insects, land snails, birds, and many reptiles excrete uric acid.

Everyday Connection

Gout

Mammals use uric acid crystals as an **antioxidant** in their cells. However, too much uric acid tends to form kidney stones and may also cause a painful condition called gout, where uric acid crystals accumulate in the joints, as illustrated in [link]. Food choices that reduce the amount of nitrogenous bases in the diet help reduce the risk of gout. For example, tea, coffee, and chocolate have purine-like compounds, called xanthines, and should be avoided by people with gout and kidney stones.

Gout causes the inflammation visible in this person's left big toe joint. (credit: "Gonzosft"/ Wikimedia Commons)

Section Summary

Ammonia is the waste produced by metabolism of nitrogen-containing compounds like proteins and nucleic acids. While aquatic animals can easily excrete ammonia into their watery surroundings, terrestrial animals have evolved special mechanisms to eliminate the toxic ammonia from their systems. Urea is the major byproduct of ammonia metabolism in vertebrate animals. Uric acid is the major byproduct of ammonia metabolism in birds, terrestrial arthropods, and reptiles.

Review Questions

BUN is ______.

1. blood urea nitrogen
2. blood uric acid nitrogen
3. an indicator of blood volume
4. an indicator of blood pressure

A

Human beings accumulate ______ before excreting nitrogenous waste.

1. nitrogen
2. ammonia
3. urea
4. uric acid

C

Critical Thinking Questions

In terms of evolution, why might the urea cycle have evolved in organisms?

It is believed that the urea cycle evolved to adapt to a changing environment when terrestrial life forms evolved. Arid conditions probably led to the evolution of the uric acid pathway as a means of conserving water.

Compare and contrast the formation of urea and uric acid.

The urea cycle is the primary mechanism by which mammals convert ammonia to urea. Urea is made in the liver and excreted in urine. The urea cycle utilizes five intermediate steps,

catalyzed by five different enzymes, to convert ammonia to urea. Birds, reptiles, and insects, on the other hand, convert toxic ammonia to uric acid instead of urea. Conversion of ammonia to uric acid requires more energy and is much more complex than conversion of ammonia to urea.

Glossary

ammonia
: compound made of one nitrogen atom and three hydrogen atoms

ammonotelic
: describes an animal that excretes ammonia as the primary waste material

antioxidant
: agent that prevents cell destruction by reactive oxygen species

blood urea nitrogen (BUN)
: estimate of urea in the blood and an indicator of kidney function

urea cycle
: pathway by which ammonia is converted to urea

ureotelic

describes animals that secrete urea as the primary nitrogenous waste material

uric acid
: byproduct of ammonia metabolism in birds, insects, and reptiles

Hormonal Control of Osmoregulatory Functions

By the end of this section, you will be able to do the following:

- Explain how hormonal cues help the kidneys synchronize the osmotic needs of the body
- Describe how hormones like epinephrine, norepinephrine, renin-angiotensin, aldosterone, anti-diuretic hormone, and atrial natriuretic peptide help regulate waste elimination, maintain correct osmolarity, and perform other osmoregulatory functions

While the kidneys operate to maintain osmotic balance and blood pressure in the body, they also act in concert with hormones. Hormones are small molecules that act as messengers within the body. Hormones are typically secreted from one cell and travel in the bloodstream to affect a target cell in another portion of the body. Different regions of the nephron bear specialized cells that have receptors to respond to chemical messengers and hormones. [link] summarizes the hormones that control the osmoregulatory functions.

Hormones That

Affect Osmoregulation Hormone	Where produced	Function
Epinephrine and Norepinephrine	Adrenal medulla	Can decrease kidney function temporarily by vasoconstriction
Renin	Kidney nephrons	Increases blood pressure by acting on angiotensinogen
Angiotensin	Liver	Angiotensin II affects multiple processes and increases blood pressure
Aldosterone	Adrenal cortex	Prevents loss of sodium and water
Anti-diuretic hormone (vasopressin)	Hypothalamus (stored in the posterior pituitary)	Prevents water loss
Atrial natriuretic peptide	Heart atrium	Decreases blood pressure by acting as a vasodilator and increasing glomerular filtration rate;

decreases sodium reabsorption in kidneys

Epinephrine and Norepinephrine

Epinephrine and norepinephrine are released by the adrenal medulla and nervous system respectively. They are the flight/fight hormones that are released when the body is under extreme stress. During stress, much of the body's energy is used to combat imminent danger. Kidney function is halted temporarily by epinephrine and norepinephrine. These hormones function by acting directly on the smooth muscles of blood vessels to constrict them. Once the afferent arterioles are constricted, blood flow into the nephrons stops. These hormones go one step further and trigger the **renin-angiotensin-aldosterone** system.

The renin-angiotensin-aldosterone system increases blood pressure and volume. The hormone ANP has antagonistic effects. (credit: modification of work by Mikael Häggström)

Renin-Angiotensin-Aldosterone

The renin-angiotensin-aldosterone system, illustrated in [link] proceeds through several steps

to produce **angiotensin II**, which acts to stabilize blood pressure and volume. Renin (secreted by a part of the juxtaglomerular complex) is produced by the granular cells of the afferent and efferent arterioles. Thus, the kidneys control blood pressure and volume directly. Renin acts on angiotensinogen, which is made in the liver and converts it to **angiotensin I**. **Angiotensin converting enzyme (ACE)** converts angiotensin I to angiotensin II. Angiotensin II raises blood pressure by constricting blood vessels. It also triggers the release of the mineralocorticoid aldosterone from the adrenal cortex, which in turn stimulates the renal tubules to reabsorb more sodium. Angiotensin II also triggers the release of **anti-diuretic hormone (ADH)** from the hypothalamus, leading to water retention in the kidneys. It acts directly on the nephrons and decreases glomerular filtration rate. Medically, blood pressure can be controlled by drugs that inhibit ACE (called ACE inhibitors).

[missing_resource: Figure_41_05_01.jpg]

Mineralocorticoids

Mineralocorticoids are hormones synthesized by the adrenal cortex that affect osmotic balance. Aldosterone is a mineralocorticoid that regulates sodium levels in the blood. Almost all of the sodium in the blood is reclaimed by the renal tubules under the influence of aldosterone. Because sodium is

always reabsorbed by active transport and water follows sodium to maintain osmotic balance, aldosterone manages not only sodium levels but also the water levels in body fluids. In contrast, the aldosterone also stimulates potassium secretion concurrently with sodium reabsorption. In contrast, absence of aldosterone means that no sodium gets reabsorbed in the renal tubules and all of it gets excreted in the urine. In addition, the daily dietary potassium load is not secreted and the retention of K^+ can cause a dangerous increase in plasma K^+ concentration. Patients who have Addison's disease have a failing adrenal cortex and cannot produce aldosterone. They lose sodium in their urine constantly, and if the supply is not replenished, the consequences can be fatal.

Antidiuretic Hormone

As previously discussed, antidiuretic hormone or ADH (also called **vasopressin**), as the name suggests, helps the body conserve water when body fluid volume, especially that of blood, is low. It is formed by the hypothalamus and is stored and released from the posterior pituitary. It acts by inserting aquaporins in the collecting ducts and promotes reabsorption of water. ADH also acts as a vasoconstrictor and increases blood pressure during hemorrhaging.

Atrial Natriuretic Peptide Hormone

The atrial natriuretic peptide (ANP) lowers blood pressure by acting as a **vasodilator**. It is released by cells in the atrium of the heart in response to high blood pressure and in patients with sleep apnea. ANP affects salt release, and because water passively follows salt to maintain osmotic balance, it also has a diuretic effect. ANP also prevents sodium reabsorption by the renal tubules, decreasing water reabsorption (thus acting as a diuretic) and lowering blood pressure. Its actions suppress the actions of aldosterone, ADH, and renin.

Section Summary

Hormonal cues help the kidneys synchronize the osmotic needs of the body. Hormones like epinephrine, norepinephrine, renin-angiotensin, aldosterone, anti-diuretic hormone, and atrial natriuretic peptide help regulate the needs of the body as well as the communication between the different organ systems.

Review Questions

Renin is made by ______.

1. granular cells of the juxtaglomerular apparatus
2. the kidneys
3. the nephrons
4. all of the above

A

Patients with Addison's disease ______.

1. retain water
2. retain salts
3. lose salts and water
4. have too much aldosterone

C

Which hormone elicits the “fight or flight” response?

1. epinephrine
2. mineralcorticoids
3. anti-diuretic hormone
4. thyroxine

A

Critical Thinking Questions

Describe how hormones regulate blood pressure, blood volume, and kidney function.

Hormones are small molecules that act as messengers within the body. Different regions of the nephron bear specialized cells, which have receptors to respond to chemical messengers and hormones. The hormones carry messages to the kidney. These hormonal cues help the kidneys synchronize the osmotic needs of the body. Hormones like epinephrine, norepinephrine, renin-angiotensin, aldosterone, anti-diuretic hormone, and atrial natriuretic peptide help regulate the needs of the body as well as the communication between the different organ systems.

How does the renin-angiotensin-aldosterone mechanism function? Why is it controlled by the kidneys?

The renin-angiotensin-aldosterone system acts through several steps to produce angiotensin II, which acts to stabilize blood pressure and volume. Thus, the kidneys control blood

pressure and volume directly. Renin acts on angiotensinogen, which is made in the liver and converts it to angiotensin I. ACE (angiotensin converting enzyme) converts angiotensin I to angiotensin II. Angiotensin II raises blood pressure by constricting blood vessels. It triggers the release of aldosterone from the adrenal cortex, which in turn stimulates the renal tubules to reabsorb more sodium. Angiotensin II also triggers the release of anti-diuretic hormone from the hypothalamus, which leads to water retention. It acts directly on the nephrons and decreases GFR.

Glossary

angiotensin converting enzyme (ACE)
: enzyme that converts angiotensin I to angiotensin II

angiotensin I
: product in the renin-angiotensin-aldosterone pathway

angiotensin II
: molecule that affects different organs to increase blood pressure

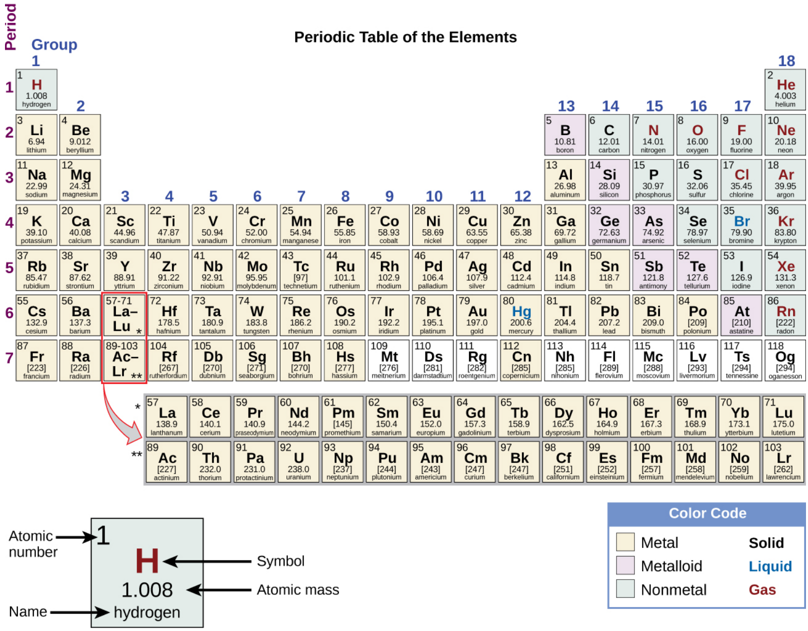
anti-diuretic hormone (ADH)
: hormone that prevents the loss of water

renin-angiotensin-aldosterone
: biochemical pathway that activates angiotensin II, which increases blood pressure

vasodilator
: compound that increases the diameter of blood vessels

vasopressin
: another name for anti-diuretic hormone

The Periodic Table of Elements



Measurements and the Metric System

Measurement	Unit	Abbreviation	Metric Equivalent	Approximate Standard Equivalent
Length	nanometer	nm	1 nm = 10^{-9} m	1 mm = 0.039 inch 1 cm = 0.394 inch 1 m = 39.37 inches 1 m = 3.28 feet 1 m = 1.093 yards 1 km = 0.621 miles
	micrometer	μm	1 μm = 10^{-6} m	
	millimeter	mm	1 mm = 0.001 m	
	centimeter	cm	1 cm = 0.01 m	
	meter	m	1 m = 100 cm 1 m = 1000 mm	
	kilometer	km	1 km = 1000 m	
Mass	microgram	μg	1 μg =	1 g = 0.035 ounce

			10^{-6} g	1 kg = 2.205 pounds
	milligram	mg	1 mg = 10^{-3} g	
	gram	g	1 g = 1000 mg	
	kilogram	kg	1 kg = 1000 g	
Volume	microliter	µl	1 µl = 10^{-6} l	1 ml = 0.034 fluid ounce 1 l = 1.057 quarts 1 kl = 264.172 gallons
	milliliter	ml	1 ml = 10^{-3} l	
	liter	l	1 l = 1000 ml	
	kiloliter	kl	1 kl = 1000 l	
Area	square centimeter	cm^2	1 cm^2 = 100 mm^2	1 cm^2 = 0.155 square inch 1 m^2 = 10.764 square feet 1 m^2 = 1.196 square yards 1 ha = 2.471 acres
	square meter	m^2	1 m^2 = 10,000 cm^2	
	hectare	ha	1 ha = 10,000 m^2	
Temperature	Celsius	°C	—	1 °C = 5/9 × (°F − 32)